EVALUATION OF OBSERVATION NETWORK BENEFITS ON SHORT-TERM WEATHER FORECASTS FOR ENERGY-METEOROLOGY APPLICATIONS

I N A U G U R A L - D I S S E R T A T I O N
ZUR
ERLANGUNG DES DOKTORGRADES
DER MATHEMATISCH-NATURWISSENSCHAFTLICHEN FAKULTÄT
DER UNIVERSITÄT ZU KÖLN

vorgelegt von

Paraskevi Vourlioti

aus Samos

Cologne, 2020

Berichterstatter:	PD Dr. H. Elbern
	Prof. Dr. S. Crewell

Tag der mündlichen Prüfung: 03. Mai 2019

Abstract

The wind and solar energy sector relies on skillful Numerical Weather Prediction (NWPs) models to operate the power transmission grid resiliently. Moreover, accurate weather predictions are important for scheduling industrial power plants, but also for price fixing the energy stock market. One of the decisive factors for forecast skills is the quality of the initial state of the model, which may help to avoid or mitigate gross prediction failures repeatedly occurring during specific weather situations. Through data assimilation, the best estimate of the atmospheric state is given by a combination of the forecasting model with observations, where the selection of measurements by type, location, and time in terms of their information value is essential. A monitoring tool, capable to quantitatively measure the capabilities of the observation network on the short-term forecast error is implemented and evaluated in this study for energy-meteorology applications.

The focus is placed on cases with exceptionally large prediction errors of solar and wind power, that challenge the weather centers and the Transmission System Operators (TSOs). To fulfill this request, a fortnight 01.-15.08.2014 that includes an exceptionally large error event on 09.08.2014 was selected. A satellite versus ground-based observation network configuration was chosen and a ranking list of the most beneficial observations was calculated. The Infrared Atmospheric Sounding Interferometer (IASI) onboard the Meteorological Operational Satellite (MetOp), as a precursor of a future sensor in a geostationary orbit, was evaluated against the classical Surface Synoptic Observations (SYNOP).

In terms of number of radiance channels per pixel, the bulk IASI data were found to contribute twice as much to the value of the forecasts in the case study compared to the SYNOP observations. Despite this, the normalized by number of observations results showed the SYNOP observations to be qualitatively superior. Especially the assimilation of the wind components were dominating the intraday forecast error reduction. The introduction in the assimilation of nine IASI water vapor channels added value to the model's forecasts, in contrast to synoptic humidity observations. The evaluation of the impact results per synoptic station location and IASI channel revealed both the most and least beneficial stations and channels.

With this method potentially all the observations for which data assimilation is prepared to assimilate, can be evaluated and models can be configured with the most favorable observations.

Kurzzusammenfassung

Die Wind und Solarenergieindustrie ist auf verlässliche Wettervorhersagemodelle angewiesen, um Übertragungsnetze sicher zu betreiben. Ferner sind genaue Wettervorhersagen wichtig, um Kraftwerkseinsätze effizient zu planen und am Energiemarkt realistische Preisbildung zu gewährleisten. Eine der entscheidenden Vorbedingungen guter Wettervorhersagen ist die Güte der Anfangswerte des Modells, die dazu beiträgt, Fälle schwerer Fehlvorhersagen zu vermeiden, die bei schwierigen Wetterlagen auftreten können. Durch Datenassimilation ist eine optimale Schätzung des atmosphärischen Zustandes möglich, indem eine geeignete Kombination des Modells mit Beobachtungen vorgenommen wird. Das Verfahren, der Ort und die Zeit der Beobachtungen sind dabei wesentlich für die Güte.

In dieser Studie wird ein Auswerteverfahren eingerichtet und angewandt, welches in der Lage ist, die Güte von Beobachtungsnetzwerken und ihrer Messungen für energiemeteorologische Anwendungen quantitativ zu bewerten. Der Schwerpunkt liegt hierbei auf einer Analyse von größeren Fehlvorhersagen von Solar- und Windenergie, die Übertragungsnetzbetreiber aus einer Kombination von Wettervorhersagen ermittelt hatten. Dabei wurde eine zweiwöchige Episode vom 01. bis 15.8.2014 ausgewertet, die eine außergewöhnlich grobe Fehlvorhersage am 9.8.2014 enthält. Hierzu wurde eine Kombination von Satellitendaten mit Bodenstationen ausgewertet und eine Bewertung der Beobachtungstypen nach ihrem Wert für eine Vorhersageverbesserung durchgeführt. Das Infrared Atmospheric Sounding Interferometer (IASI) auf die Meteorologischer Operationssatellit (MetOp) wurde gegen klassische Bodenbeobachtungen evaluiert.

Es wurde festgestellt, dass die IASI-Massendaten in Bezug auf die Anzahl der Strahlungskanäle pro Pixel doppelt so viel zum Wert der Prognosen in der Fallstudie beitragen wie die SYNOP-Beobachtungen. Trotzdem zeigten die nach Anzahl der Beobachtungen normalisierten Ergebnisse, dass die SYNOP-Beobachtungen qualitativ überlegen waren. Insbesondere erwies sich die Assimilation von Windkomponentendaten als besonders wertvoll, die Vorhersagefehler innerhalb einer Tagesvorhersage zu reduzieren. Die zusätzliche Nutzung von neun IASI-Wasserdampfkanälen für die Assimilation erwies sich den synoptischen Feuchtebeobachtungen überlegen.

Die Auswertung der Beobachtungswirkung individueller Messungen oder Kanäle kann mit dem hier entwickelten Verfahren besonders günstige oder ungünstige Beobachtungskonfigurationen ermitteln, und zu möglichst effizienten Beobachtungsnetzen insbesondere auch für energiemeteorologische Awendungen führen.

Contents

List of Figures iii

List of Tables vii

Acronyms ix

1 Introduction 1

2 Formulation of Forecast Sensitivity to Observations 5

3 Weather and Research Forecasting Model 11
3.1 Variational data assimilation in WRF model 13

4 Experiment set-up 19
4.1 Observation configuration 22
4.1.1 Infrared Atmospheric Sounding Interferometer 23
4.1.2 Surface Synoptic Observations 31
4.2 Assimilation of observations 32
4.2.1 IASI assimilation 37
4.2.2 SYNOP assimilation 40

5 Observation impact results 43
5.1 Non-linear and linear forecast errors 43
5.2 IASI and SYNOP observation impact 46
5.2.1 Observation System Experiments 52
5.3 Poor predictability cases 54
5.3.1 Weather conditions 59
5.4 Evaluation of the results 64
5.4.1 Assimilation results 64
5.4.2 Impact results 69
5.4.3 Examination of observation impact 83

6 Conclusions 89

A Appendix 93
A.1 FSO configuration at 18 UTC 93

A.2 Spatial distribution of observation impact for SYNOP wind measurements 97
A.3 Model namelist configuration 102

Bibliography **111**

Acknowledgments **117**

List of Figures

2.1 Schematic of FSO 5

3.1 Horizontal and vertical model grid configuration 12
3.2 Vertical model levels 12
3.3 WPS work-flow 13
3.4 Flow-chart of WRFDA configuration 14
3.5 Observation ingestion in WRFDA 15

4.1 The cycling mode of the experiment set-up. 19
4.2 WRF domain used in the study. The terrain height is depicted in meters (m). 20
4.3 IASI on-board METOP and the first IASI/METOP-B spectrum on 24.10.2012 24
4.4 The IASI spectral channels 27
4.5 IASI observation errors 28
4.6 Emission weighting functions of nine water vapor channels of IASI . . 28
4.7 Emission weighting functions of IASI spectral channels 30
4.8 SYNOP observations 31
4.9 Assimilated observations types for the period 01.-15.08.2014 as a percentage of the total number of observations, at each assimilation hour 33
4.10 Spatial domain coverage by IASI data 35
4.11 Spatial coverage by SYNOP data 36
4.12 Bias correction scatter plots 38
4.13 Percentage of IASI data that successfully passed quality control . . . 39
4.14 Percentage of SYNOP data that successfully passed quality control . 40
4.15 Mean SYNOP observation error for 01.-15.08.2018 41

5.1 Time series of the non-linear and the approximated linear error for 01.-15.08.2014. 44
5.2 Time series of linear versus non-linear forecast errors 45
5.3 Time averaged error reduction by IASI and SYNOP 46
5.4 Number of IASI and SYNOP data assimilated in the fortnight 47
5.5 Average error reduction, normalized by observation number 48
5.7 Relative forecast error reduction 48

5.8 Time series of forecast error reduction by IASI and SYNOP, normalized by observation number 49
5.9 Spatial distribution of SYNOP stations measuring pressure that had the worst performance in the fortnight 50
5.10 Spatial distribution and channel ranking list of the least beneficial IASI assimilation in the fortnight 51
5.11 Average forecast error calculated by OSEs 53
5.12 Time series (01.-15.08.2014) of day-ahead forecasts and actual total power values 54
5.13 Time series (01.-15.08.2014) of day-ahead forecasts and actual total power values at the four assimilation hours 55
5.14 Time series of RMSE and MAE 56
5.15 RMSE per hour (00, 06, 12, 18 UTC) for 01.-15.082014. 56
5.16 Time series of days-ahead forecasts and actual total power values for wind power, on days of large error events 57
5.17 Time series of days-ahead forecasts and actual total power values for wind power, on days of large error events. 57
5.18 Geopotential height in decametres at 500 hPa and sea level pressure plots, for the poor predictability cases 59
5.19 Surface analyses by DWD and SEVIRI imagery on days of poor predictability cases(a) 61
5.20 Surface analyses by DWD and SEVIRI imagery on days of poor predictability cases(b) 62
5.21 Time series of innovation and analysis residuals, for assimilation of SYNOP data 66
5.22 Time series of innovation and analysis residuals, for assimilation of IASI data(a) 67
5.23 Time series of innovation and analysis residuals, for assimilation of IASI data(b) 68
5.24 Observation impact per assimilation hour for the European domain . 70
5.25 Relative forecast error reduction by assimilation of IASI and SYNOP 71
5.26 Relative forecast error reduction by assimilation of SYNOP measurements and IASI spectral channels 72
5.27 Spatial distribution and channel ranking list of the most beneficial SYNOP and IASI observations on 09.08, 06 UTC 74
5.28 Spatial distribution and channel ranking list of the most beneficial SYNOP and IASI observations 11.08., 06 UTC 75
5.29 Spatial distribution and channel ranking list of the most beneficial SYNOP and IASI observations on 03.08., 12 UTC 77
5.30 Spatial distribution and channel ranking list of the most beneficial SYNOP and IASI observations on 04.08.,12 UTC 78
5.31 Spatial distribution and channel ranking list of the most beneficial SYNOP and IASI observations on 09.08., 12 UTC 79
5.32 Spatial distribution and channel ranking list of the most beneficial SYNOP and IASI observations on 10.08., 12 UTC 80

5.33 Spatial distribution and channel ranking list of the most beneficial SYNOP and IASI observations on 11.08., 12 UTC 81
5.34 Spatial distribution and channel ranking list of the most beneficial SYNOP and IASI observations on 13.08., 12 UTC 82
5.35 Scatter plots of observation impact of SYNOP variables and their corresponding innovation vector. 84
5.36 Scatter plots of observation impact of IASI spectral channels and their corresponding innovation vector. 86

A.1 FSO configuration at 18 UTC 94
A.2 FSO configuration at 12 UTC 95
A.3 Impact of SYNOP stations measuring wind (assimilation of meridional V) on the 6-hour forecast, for assimilation hour 06 UTC and for all poor predictability cases. 98
A.4 Impact of SYNOP stations measuring wind (assimilation of meridional V) on the 6-hour forecast, for assimilation hour 12 UTC and for all poor predictability cases. 99
A.5 Impact of SYNOP stations measuring wind (assimilation of zonal U) on the 6-hour forecast, for assimilation hour 06 UTC and for all poor predictability cases. 100
A.6 Impact of SYNOP stations measuring wind (assimilation of zonal U) on the 6-hour forecast, for assimilation hour 12 UTC and for all poor predictability cases. 101

List of Tables

2.1 Forecast error approximations 9

4.1 WRF model configuration 20
4.2 Computational resources on JURECA. One FSO cycle demands the execution of WRF forward model, assimilation and adjoint model. The calculation of the impact is done with the help of the WRF assimilation model. 21
4.3 Spectral range given in units of cm^{-1} and μm along with the corresponding IASI application (Chalon et al. [2001]). 25
4.4 Observation types assimilated to produce the reference analysis in the FSO scheme 32

5.1 Dates with highest RMSE and MAE scores for wind and solar power. 56
5.2 Relative forecast error in percentage for days with the largest RMSE and MAE scores. 58

A.1 WRF namelist for the forward runs. 103
A.2 WRFVAR namelist for the assimilation runs. 104
A.3 WRFPLUS namelist for the adjoint runs. 106
A.4 WRFVAR namelist for impact calculation runs. 107

Acronyms

AFWA	Air Force Weather Agency
AIRS	Atmospheric Infrared Sounder
ALADIN	Aire Limitée Adaptation dynamique Développement InterNational
AMSU-A	Advanced Microwave Sounding Unit-A
ARPEGE	Action de Recherche Petite Echelle Grande Echelle
ARW	Advanced Research WRF
ASCAT	Advanced Scatterometer
ATOVS	Advanced TIROS Operational Vertical Sounder
AVHRR	Advanced Very High Resolution Radiometer
CG	Conjugate Gradient
CRTM	Community Radiative Transfer Model
DWD	Deutscher WetterDienst
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
ECMWF	European Centre for Medium-Range Weather Forecasts
EFOV	Effective Field Of View
EPS	EUMETSAT Polar System
FAA	Federal Aviation Administration
FASTEX	Fronts and Atlantic Storm-Track EXperiment
FEC	Forecast Error Contribution
FOV	Field Of View
FSL	Forecast Systems Laboratory
FSO	Forecast Sensitivity to Observations
GDAS	Global Data Assimilation System
GFS	Global Forecasting System
GMAO	Global Modeling and Assimilation Office
GNSS	Global Navigation Satellite System
GOES	Geostationary Operational Environmental Satellite
GRAS	Global Receiver for Atmospheric Sounding
GTS	Global Telecommunications System

IASI	Infrared Atmospheric Sounding Interferomenter
IR	Infrared
IRS	Infrared Sounder
JCSDA	Joint Center for Satellite Data Assimilation
JSC	Jülich Supercomputing Center
JURECA	Jülich Research on Exascale Cluster Architectures
MAE	Mean Absolute Error
METAR	Meteorological Aerodrome Report
MetOp	Meteorological Operational Satellite
MPI	Message Passing Interface
MSG	MeteoSat Second Generation
MTG	MeteoSat Third Generation
NAE	North Atlantic and European model
NASA	National Aeronautics and Space Administration
NAVDAS	NRL Atmospheric Variational Data Assimilation
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NESDIS	National Environmental Satellite Data and Information Service
NMC	National Meteorological Center
NOGAPS	Nany Operational Global Atmospheric Prediction System
NRL	Naval Research Laboratory
NWP	Numerical Weather Prediction
OMA	Observation Minus Analysis
OMB	Observation Minus Background
OSE	Observation System Experiment
OSSE	Observation System Simulation Experiments
PREPBUFR	Prepared Binary Universal Form for the Representation of meteorological data
RMSE	Root Mean Square Error
RTE	Radiative Transfer Equation
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SFR	Spectral Response Function
SV	Singular Vectors
SSM/I	Special Sensor Microwave Image
SYNOP	Surface Synoptic Observations
TSO	Transmission System Operators
TCW	Total Cloud Water
VarBC	Variational Bias Correction

WMO	World Meteorological Organization
WPS	WRF Preprocessing System
WRF	Weather amd Research Forecast
WRFDA	WRF Data Assimilation System
3DVAR	3 Dimentional Variational Assimilation
4DVAR	4 Dimentional Variational Assimilation

Chapter 1

Introduction

With a target of supplying 27% share of energy consumption from renewable sources by 2030 (European Parliament [2014]), the research field of energy meteorology, comprising wind and solar energy, is flourishing. Accurate short-term wind and solar power forecasts are vital for maintaining power grid stability, scheduling the power plants but also for the energy stock market (Kleissl [2013]). The day-ahead solar and wind power forecasts are typically provided by NWPs, which simulate the physical processes of the atmosphere based on the primitive equations.

The key to the accurate day-ahead solar and wind power forecasts lies in the performance of the NWPs. Some of the main drivers of the performance are the measurements availability for the initial state, the physical parameterizations and the computational resources to compute with high spatial resolutions (Kleissl [2013]). Finding the optimal way to initialize a model brings into focus the research field of data assimilation. Data assimilation is the technique that combines observations and NWPs along with their corresponding errors to provide the best estimate of the atmospheric state, called the analysis (Kalnay [2003]). The analysis can be used in various ways, including initialization, studying the climate through reanalyses, examining individual components of the existing observation network by conducting Observation System Experiments (OSEs), and predicting the potential impact of new components of the observation network via Observation System Simulation Experiments (OSSEs) (Barker et al. [2003]).

Energy meteorology and data assimilation can be combined in an fruitful way for optimal observation network configuration, in terms of wind and solar power generated energy prediction. In particular, the information gain by the different types of assimilated observations can be assessed and monitored with the help of well defined properties of the assimilation system. The relative impact of the observation network on the short-term forecast can then be investigated in terms of spatio-temporal configuration.

In the frame of data assimilation, the impact of observation data sets on the quality of the forecast is traditionally estimated by the OSEs. These experiments demand the removal of subsets of observations from a data assimilation system and the resulting forecast is compared against a control set that includes all observations (Daley [1993]). But, because of their computational expense, they usually involve

only a small number of independent experiments (Gelaro and Zhu [2009]). In recent years a new algorithm relying on the adjoint (i.e. mathematical transpose) properties of the data assimilation system has been developed and measures the response of the forecast error to all perturbations, including all available observations in the experiment. This algorithm does not only include all the observations in the system, leaving the gain matrix untouched (Cardinali [2009]) but is also computationally cheap, allowing a monitoring aspect of the performance of the observation network relative to weather conditions. This advanced and sophisticated tool for estimating the data impact on the forecasting system is called Forecast Sensitivity to Observations (FSO)(Langland and Baker [2004]). FSO answers the question where and which observations should be placed (by the sensitivity gradients) and how much each observation (by variable or by instrument) has contributed to improve the forecast. The latter can be viewed as complementing the OSEs and the answer to the former falls in the targeted observations research.

The concept of targeted observations addresses the question of identifying in advance, regions where additional observations would improve the forecast. For instance, Langland et al. [1999] assessed the impact of observations placed in real time in regions identified apriori, as target regions, by the most rapidly growing singular vectors (SVs) of the linearized model in the context of the Fronts and Atlantic Storm-Track EXperiment (FASTEX) during January and February of 1997. They employed dropsonde and satellite wind data on the sensitive locations and demonstrated improved 24 hour forecast skill with this technique. More specific, the SV method defines a matrix problem, consisting of the tangent and the adjoint model (transpose of the linearized forecast model) along with a scaling matrix, which is usually an energy norm (Baker and Daley [2000]). The SVs of the adjoint model maximize the growth of the perturbation energy. Studies of Gelaro et al. [1998] and Palmer et al. [1998] supported the applicability of the total energy norm for studying the atmospheric predictability within the time scales of NWPs. They demonstrated this by showing that the spectra of singular values are dominant in the wavenumber band for which the analysis error variance is relative small (Palmer et al. [1998]). The SVs approach for the targeted observations is based on the fact that corrections to initial conditions of the forecasts, result in improved forecast skills (Rabier et al. [1996]). The success of the SVs is attributed to the fraction of the analysis error that projects into the leading SVs. The growth of this component of the analysis error can dominate the forecast error (Gelaro et al. [1999]). Both the SV and the adjoint sensitivity method are related in a mathematical sense. Analyzing the FASTEX campaign they pinpointed to similar sensitive areas.

Following the introduction of the FSO tool, many weather centers have extended their data assimilation systems by FSO to estimate the distinct value of various data sets on the short term forecast. For example, the Naval Research Laboratory and National Aeronautics and Space Administration/Global Modeling and assimilation Office (NASA/GMAO) have adopted the scheme on an operational basis to monitor the observation network assimilated in their models (Kalnay et al. [2012]).

Langland and Baker [2004], calculated the 24 hour global forecast error of the Navy Operational Global Atmospheric Prediction System (NOGAPS) during June and December 2002. They used the Naval Research Laboratory (NRL) Atmospheric Variational Data Assimilation (NAVDAS) system to perform the 3 Dimensional Variational Assimilation (3DVAR) of space borne observations (e.g. from the Geostationary Operational Environmental Satellite (GOES) and the Advanced TIROS Operational Vertical Sounder (ATOVS)) as well as surface data from ships, buoys and other in-situ devices. It was found that 60% of the global error reduction was due to observations below 500 hPa and they revealed a strong correlation between observation impact and cloud cover at observation location. For instance, in locations with 80 % cloud cover the impact of rawinsonde temperature observation was found to be 1,5 times larger than in locations with 20 % cloud cover. Two explanations were given for this finding, the first one referring to the study of McNally [2002] claiming cloudy regions to be more sensitive to analysis errors and the second one arguing that in cloudy regions fewer observation are available for assimilation (e.q due to quality control) and thus are given more weight in the assimilation procedure. These findings support the concept of adding information (i.e observations) into cloudy regions could be a means to improve the forecast skill.

Joo et al. [2013], studied the value of space borne observations on the 24 hour forecast using the FSO scheme, adapted for the Met Office global NWP model and the 4 Dimensional Variational Assimilation (4DVAR) assimilation system. The experiment was conducted from 22.08. to 29.09 2010. Compared to ground based observations, satellite data were found to account for 64 % of the forecast error reduction. Moreover microwave and hyperspectral infrared soundings had the largest total impacts. For satellite platforms, MetOp-A had the largest impact. This was attributed to four sensors: IASI (Infrared Atmospheric Sounding Interferometer), GNSS (Global Navigation Satellite System), Global Receiver for Atmospheric Sounding (GRAS) and the Advanced Scatterometer (ASCAT). A major source of forecast improvement from a ground based observations type were the assimilated Meteorological Aerodrome Reports (METAR) (Joo et al. [2013]). The study did not account for varying atmospheric conditions but concluded that FSO can be exploited in this manner.

Jung et al. [2013], applied the FSO scheme in East Asia and the western Pacific for the 2008 typhoon season (16.08. - 01.10.2008). They were the first to examine the capability of the scheme in the limited area Weather Research and Forecasting (WRF) model and the sensitivity to the error covariance parameters. They used the 3DVAR assimilation system with a 45 kilometer resolution, and calculated the 6-hour forecast impact from different types of observations including ground based synoptic reports and satellite observations like the Advanced Microwave Sounding Unit-A (AMSU-A), but no infrared observations. They discovered that for this application the fraction of beneficial observations was higher (60 % -70 % of observation were beneficial) than reported in previous studies and that in the ranking list, the synoptic reports were the most beneficial after the satellite AMSU-A observations. The impact of satellite measured temperature brightness was found to be one order of magnitude greater than the one from conventional observations, as the magnitude is also proportional

to the observation number. As for the conventional observations, in agreement with the World Meteorological Organization (WMO) Andersson [2008], the impact of momentum variables was found to be greater than those of moisture and mass. Their results were also evaluated successfully with OSEs.

This study aims to apply and evaluate the FSO algorithm for the specific needs of energy meteorology, where skillful short-term forecasts are vital. Emphasis should be placed for episodes of exceptionally poor predictability. That is for cases where even weather centers collectively produced erroneous forecasts, which resulted in grid management difficulties and ruined power trading revenues. A ranking list of the most valuable assimilated observations for the short-term forecast should be identified, indicating the values of observation types and individual measurements. For the satellite data, particular attention should be given to sensors that are expected to play a leading role in the future. This request leads to the evaluation of IASI, onboard MetOp, serving as a precursor to the Infrared Sounder (IRS) on MeteoSat Third Generation (MTG) (EUMETSAT [2019]). The classical in-situ synoptic reports, should be evaluated against the prominent satellite data.

The formulation of the algorithm is described in Chapter 2. In Chapter 3, the NWP model and the radiative transfer model utilized for this study are presented. The configuration of the experiment, including experiment set-up and information on the examined observations are given in Chapter 4. Next, the aggregated results for the examined time span, the OSE experiments and the focus on low solar and wind power predictability days are given in Chapter 5. The conclusions and discussion of the results are presented in Chapter 6.

Chapter 2

Formulation of Forecast Sensitivity to Observations

Forecast Sensitivity to Observations (FSO) is an adjoint based technique that allows the calculation of the observations impact on the assimilation result, commonly termed analysis, at each distinct assimilation cycle. Subsequently, an analysis as defined in section (Section 3.1) is calculated first. The goal of FSO is then to quantitatively measure the forecast error reduction caused by each distinct observation or observation type. The configuration of FSO is depicted on Fig. 2.1, where a 3DVAR analysis is calculated at time $\mathbf{t_0}$. In the next paragraphs, FSO will be formulated (Langland and Baker [2004]) by unraveling this configuration, to end up with the core of the FSO equation.

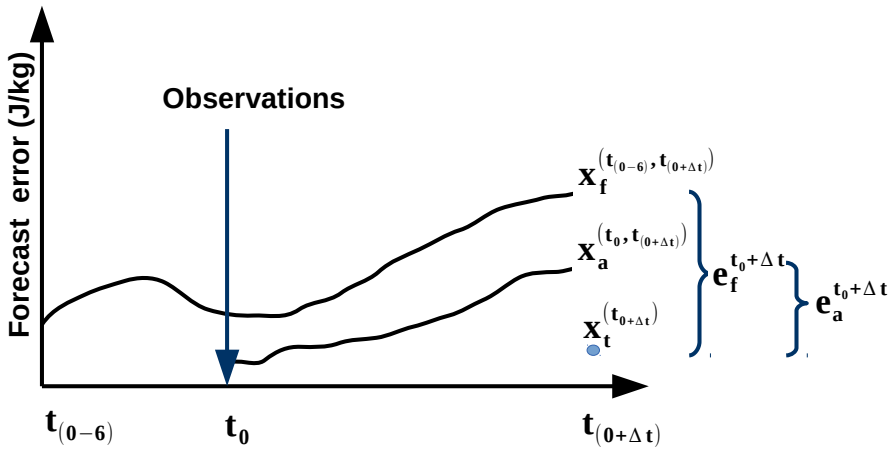


Figure 2.1: Schematic of FSO. Two forecasts, $({x_f}^{(t_{(0-6)},t_{(0+\Delta t)})}, {x_a}^{(t_0,t_{(0+\Delta t)})})$, are initiated with a difference of six hours and evaluated at verification time after six hours $(t_{0+\Delta t})$. The error measure is defined upon verification time as the difference of the two forecasts from the reference analysis ${x_t}^{(t_{0+\Delta t})}$.

After the assimilation is performed at $\mathbf{t_0}$, two forecasts are initiated: one starting from the analysis time and an other one from a model run initiated at $\mathbf{t_{0-6}}$. Both end at $\mathbf{t_{0+\Delta t}}$. The forecast starting at $\mathbf{t_{0-6}}$, serves as background information for the assimilation at $\mathbf{t_0}$. This means that at $\mathbf{t_{0+\Delta t}}$, the only difference between the two forecasts is the information gained from the observations, which are included in the forecast initiated from the analysis.

At this time step ($\mathbf{t_{0+\Delta t}}$), an overall scalar forecast error e has to be defined for which the impact of observations will be calculated. The dry total energy norm, in units of J/Kg,

$$\mathbf{e} = \frac{\mathbf{1}}{\mathbf{2}} \iiint_{\mathbf{\Sigma}} [\mathbf{u'^2} + \mathbf{v'^2} + (\frac{\mathbf{g}}{\overline{\mathbf{N}}\ \overline{\mathbf{\Theta}}})^{\mathbf{2}} \mathbf{\Theta'^2} + (\frac{\mathbf{1}}{\overline{\rho} \mathbf{c_s}})^{\mathbf{2}} \mathbf{p'^2}]\, \mathbf{d\Sigma} \tag{2.1}$$

is a suitable and commonly used choice as it consists of the most relevant model variables (namely wind speed, temperature and pressure) (Rabier et al. [1996]). The energy norm is used to calculate the forecast error cost function or forecast response. It allows the comparison of different types of measurements and physical parameters as wind and temperature by one scale. The first part of (Eq. (2.1)) specifies the kinetic energy and the second the available potential energy. More specific, u', v', Θ', p' are perturbations of zonal wind, meridional wind, potential temperature and pressure, respectively. The coefficients are, $\overline{\Theta}$ which is the potential temperature, $\overline{\rho}$ is the density, c_s is the speed of sound and $\overline{N}$ denotes the Brunt-Vaisala frequency. The $\mathbf{d\Sigma}$ denotes the integration over the domain's three dimensions $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$.

Then, two cost functions $\mathbf{J_a^{t_{0+\Delta t}}}$ and $\mathbf{J_f^{t_{0+\Delta t}}}$ that measure the contribution to the forecast errors $\mathbf{e_a^{t_{0+\Delta t}}}$ and $\mathbf{e_f^{t_{0+\Delta t}}}$ from the two forecasts $\mathbf{x_a}^{(\mathbf{t_0}, \mathbf{t_{(0+\Delta t)}})}$, $\mathbf{x_f}^{(\mathbf{t_{(0-6)}}, \mathbf{t_{(0+\Delta t)}})}$ are defined as,

$$\mathbf{J_a^{t_{0+\Delta t}}} = \frac{\mathbf{1}}{\mathbf{2}} \mathbf{e_a^{t_{0+\Delta t}}} \tag{2.2}$$

$$\mathbf{J_f^{t_{0+\Delta t}}} = \frac{\mathbf{1}}{\mathbf{2}} \mathbf{e_f^{t_{0+\Delta t}}} \tag{2.3}$$

where,

$$\mathbf{e_a^{t_{0+\Delta t}}} = (\mathbf{x_a^{(t_0, t_{0+\Delta t})}} - \mathbf{x_t^{t_{0+\Delta t}}})^{\mathbf{T}} \mathbf{C} (\mathbf{x_a^{(t_0, t_{0+\Delta t})}} - \mathbf{x_t^{t_{0+\Delta t}}}), \tag{2.4}$$

and

$$\mathbf{e_f^{t_{0+\Delta t}}} = (\mathbf{x_f^{(t_{0-6}, t_{0+\Delta t})}} - \mathbf{x_t^{t_{0+\Delta t}}})^{\mathbf{T}} \mathbf{C} (\mathbf{x_f^{(t_{0-6}, t_{0+\Delta t})}} - \mathbf{x_t^{t_{0+\Delta t}}}). \tag{2.5}$$

The forecast errors $\mathbf{e_a^{t_{0+\Delta t}}}$ and $\mathbf{e_f^{t_{0+\Delta t}}}$ are quadratic error measures of the forecasts initiated from the analysis $\mathbf{x_a^{(t_0, t_{0+\Delta t})}}$ and the earlier forecast $\mathbf{x_f^{(t_{0-6}, t_{0+\Delta t})}}$ with respect to the reference state $\mathbf{x_t}^{(\mathbf{t_{0+\Delta t}})}$. The reference state is usually an analysis provided at the verification time $\mathbf{t_{0+\Delta_t}}$ by the same model that provides the forward runs. The superscripts in the equations refer to the corresponding times as illustrated in Fig. 2.1 and $\mathbf{C}$ is a diagonal matrix containing the weights according to the total dry energy norm.

The non-linear forecast error reduction can then be defined as the difference of the two quadratic error measures $\mathbf{\Delta e_a^f} = \mathbf{e_a} - \mathbf{e_f}$ and can be exactly calculated using (Eq. (2.4)) and (2.5). This formulation provides a simple plus/minus configuration,

where negative values would imply forecast improvement due to assimilation as the forecast from the analysis would be closer to the reference than the simple forecast. Similarly positive values of the non- linear forecast error would imply degradation of the forecast due to assimilation. In order to establish a relation between forecast error at forecast and assimilation time, the adjoint model (transposed of the linearized version of the forecast model) is utilized. An approximation of $\mathbf{\Delta e_a^f}$ is sought with the use of adjoint sensitivity gradients. To achieve this, $\mathbf{\Delta e_a^f}$ can be expressed with the use of equations (Eq. (2.2)) and (2.3). The gradients of the cost function $\partial J_a^{t_0+\Delta t}/\partial \mathbf{x_a^{t_0+\Delta t}}$ and$\partial J_f^{t_0+\Delta t}/\partial \mathbf{x_f^{t_0+\Delta t}}$ are given by:

$$\partial \mathbf{J_a^{t_0+\Delta t}}/\partial \mathbf{x_a^{t_0+\Delta t}} = \mathbf{C}(\mathbf{x_a^{(t_0,t_0+\Delta t)}} - \mathbf{x_t^{t_0+\Delta t}}) \tag{2.6}$$

$$\partial \mathbf{J_f^{t_0+\Delta t}}/\partial \mathbf{x_f^{t_0+\Delta t}} = \mathbf{C}(\mathbf{x_f^{(t_0-6,t_0+\Delta t)}} - \mathbf{x_t^{t_0+\Delta t}}) \tag{2.7}$$

These gradients serve as input information to the adjoint model, starting from time $\mathbf{t_{0+\Delta t}}$ back to the assimilation time at $\mathbf{t_0}$. Two adjoint model runs are performed, from the forecast time $\mathbf{t_{0+\Delta t}}$ to assimilation time $\mathbf{t_0}$.

The adjoint model maps the sensitivity of the forecast response with respect to the forecast control vector, $\partial \mathbf{J_a^{t_0+\Delta t}}/\partial \mathbf{x_a^{t_0+\Delta t}}$ and $\partial \mathbf{J_f^{t_0+\Delta t}}/\partial \mathbf{x_f^{t_0+\Delta t}}$, into the sensitivity of the forecast response with respect to the initial conditions, $\partial \mathbf{J_a^{t_0}}/\partial \mathbf{x_a^{t_0}}$ and $\partial \mathbf{J_f^{t_0}}/\partial \mathbf{x_f^{t_0}}$, linearized along the forecast trajectories. The resulting gradients, show locations, where changes to the initial conditions have the largest impact on the forecast error. The sensitivity of the forecast error to initial humidity conditions are obtained by the adjoint of the linearized moist physical processes as a secondary effect (Janisková and Cardinali [2017]).
The sensitivity of the forecast response to the observations, $\partial \mathbf{J}/\partial \mathbf{y^{t_0}}$, can be derived from the sensitivity of the forecast response to the initial conditions by using the rule of chain,

$$\frac{\partial \mathbf{J}}{\partial \mathbf{y^{t_0}}} = \frac{\partial \mathbf{J}}{\partial \mathbf{x_a^{t_0}}} \frac{\partial \mathbf{x_a^{t_0}}}{\partial \mathbf{y^{t_0}}} \tag{2.8}$$

where the gradient $\partial \mathbf{x_a^{t_0}}/\partial \mathbf{y^{t_0}} = \mathbf{K^T}$, provides the sensitivity of the analysis system to the observations. The calculation of the adjoint of the data assimilation system is a computationally demanding task, as the inverse of the analysis error covariance matrix, $\mathbf{A^{-1}}$, has to be approximated (Daley [1993]).

$$\mathbf{K^T} = \mathbf{R^{-1}HA^{-1}}, \tag{2.9}$$

where $\mathbf{H}$ is the observation or forward operator that maps from model to observation space and $\mathbf{R}$ is the observation error covariance matrix. The Lanczos algorithm that calculates the converged eigenvalues lying in the Krylov subspace (Van der Vorst [2003]) is used to approximate the inverse of $\mathbf{A}$. To accomplish this, the minimization algorithm is switched from Conjugate Gradient (CG) to Lanczos and it can be proven that the error associated with the variation solution is given by the inverse of the Hessian matrix (second derivative of the 3DVAR cost function):

$$\mathbf{A^{-1}} = (\frac{\partial^2 \mathbf{J}}{\partial^2 \mathbf{x}}) = \mathbf{B^{-1}} + \mathbf{H^T R^{-1} H}. \tag{2.10}$$

Eq. (2.10) states that the precision of the analysis error covariance is equal to the sum of the precision of the background information and the observations passed through the observation operator (Zou et al. [1997]). Now, Eq. (2.8) becomes:

$$\frac{\partial \mathbf{J}_\mathbf{a}^{\mathbf{t_0}}}{\partial \mathbf{y}^{\mathbf{t_0}}} = \mathbf{K}^\mathbf{T} \frac{\partial \mathbf{J}_\mathbf{a}^{\mathbf{t_0}}}{\partial \mathbf{x}_\mathbf{a}^{\mathbf{t_0}}} \tag{2.11}$$

for the sensitivity stemming from the analysis $\mathbf{x_a}$ and

$$\frac{\partial \mathbf{J}_\mathbf{f}^{\mathbf{t_0}}}{\partial \mathbf{y}^{\mathbf{t_0}}} = \mathbf{K}^\mathbf{T} \frac{\partial \mathbf{J}_\mathbf{a}^{\mathbf{t_0}}}{\partial \mathbf{x}_\mathbf{f}^{\mathbf{t_0}}} \tag{2.12}$$

for the sensitivity stemming from the forecast $\mathbf{x_f}$. Similarly to the sensitivity of the response function to the initial conditions, the sensitivity of the response function to the observations indicates locations where small changes to the observations have the greatest impact on the forecast error. Baker and Daley [2000] derived the sensitivity of the forecast response to initial conditions at time $\mathbf{t_0}$ as:

$$\frac{\partial \mathbf{J}}{\partial \mathbf{x}^{\mathbf{t_0}}} = \frac{\partial \mathbf{J}_\mathbf{a}^{\mathbf{t_0}}}{\partial \mathbf{x}_\mathbf{a}^{\mathbf{t_0}}} + \frac{\partial \mathbf{J}_\mathbf{f}^{\mathbf{t_0}}}{\partial \mathbf{x}_\mathbf{f}^{\mathbf{t_0}}} \tag{2.13}$$

The non-linear error can then be expressed in terms of gradients at forecast time.

$$\mathbf{\Delta e_a^f} = \langle (\mathbf{x_a} - \mathbf{x_f})^{(\mathbf{t_0}+\mathbf{\Delta t})}, (\frac{\partial \mathbf{J}_\mathbf{a}^{(\mathbf{t_0}+\mathbf{\Delta t})}}{\partial \mathbf{x}_\mathbf{a}^{(\mathbf{t_0},\mathbf{t_0}+\mathbf{\Delta t})}} + \frac{\partial \mathbf{J}_\mathbf{f}^{(\mathbf{t_0}+\mathbf{\Delta t})}}{\partial \mathbf{x}_\mathbf{f}^{(\mathbf{t_0}-\mathbf{6},\mathbf{t_0}+\mathbf{\Delta t})}}) \rangle \tag{2.14}$$

The connection between the forecast error at forecast time with the error at assimilation time can now be made. The approximation of $\mathbf{\Delta e_a}^\mathbf{f}$ can be written as:

$$\delta \mathbf{e_a^f} = \langle \delta \mathbf{x_a}, (\frac{\partial \mathbf{J}}{\partial \mathbf{x}_\mathbf{a}^{\mathbf{t_0}}} + \frac{\partial \mathbf{J}}{\partial \mathbf{x}_\mathbf{f}^{\mathbf{t_0}}}) \rangle \tag{2.15}$$

where $\delta \mathbf{x_a}$ is the analysis increment and $\delta \mathbf{e_a^f}$ becomes:

$$\delta \mathbf{e_a^f} = \langle \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_\mathbf{f}^{\mathbf{t_0}}), (\frac{\partial \mathbf{J}}{\partial \mathbf{x}_\mathbf{a}^{\mathbf{t_0}}} + \frac{\partial \mathbf{J}}{\partial \mathbf{x}_\mathbf{f}^{\mathbf{t_0}}}) \rangle \tag{2.16}$$

$\delta \mathbf{e_a^f}$ is not an exact calculation as $\mathbf{\Delta e_a}^\mathbf{f}$, because the sensitivity gradients are calculated with the use of the adjoint model. Using the adjoint property,

$$\delta \mathbf{e_a^f} = \langle (\mathbf{y} - \mathbf{H}\mathbf{x}_\mathbf{f}^{\mathbf{t_0}}), \mathbf{K}^\mathbf{T} (\frac{\partial \mathbf{J}}{\partial \mathbf{x}_\mathbf{a}^{\mathbf{t_0}}} + \frac{\partial \mathbf{J}}{\partial \mathbf{x}_\mathbf{f}^{\mathbf{t_0}}}) \rangle \tag{2.17}$$

the approximation can be written as

$$\delta \mathbf{e_a^f} = \langle (\mathbf{y} - \mathbf{H}\mathbf{x}_\mathbf{f}^{\mathbf{t_0}}), \frac{\partial \mathbf{J}}{\partial \mathbf{y}^{\mathbf{t_0}}} \rangle \tag{2.18}$$

The final step to quantitatively estimate the observation impact in the forecast error is to calculate the inner product of the sensitivity of the response function to

observation and the innovation vector $\mathbf{d}$ (see Eq. (3.2)). This means the greater the innovation vector, the greater the impact. The Forecast Error Contribution (FEC) by the assimilation of observations or observation impact is defined as:

$$\delta \mathbf{e_a^f} = \langle \mathbf{d}, \frac{\partial \mathbf{J}}{\partial \mathbf{y^{t_0}}} \rangle. \tag{2.19}$$

According to Gelaro et al. [2007], various orders of approximation of observation impact can be derived with the use of Taylor expansion. More specific, the acquired variations are presented in Table 2.1.

Table 2.1: The first ($\delta\mathbf{e1}$), second ($\delta\mathbf{e2}$) and third ($\delta\mathbf{e3}$) order Taylor approximations. $\delta\mathbf{e4}$ and $\delta\mathbf{e5}$ are augmented versions of the third order approximation.

Measure	Formula
$\mathbf{\Delta e}$	$(\mathbf{x_a^{(t_0,t_{0+\Delta t})}} - \mathbf{x_t^{t_{0+\Delta t}}})^{\mathbf{T}}\mathbf{C}(\mathbf{x_a^{(t_0,t_{0+\Delta t})}} - \mathbf{x_t^{t_{0+\Delta t}}}) - (\mathbf{x_f^{(t_{0-6},t_{0+\Delta t})}} - \mathbf{x_t^{t_{0+\Delta t}}})^{\mathbf{T}}\mathbf{C}(\mathbf{x_f^{(t_{0-6},t_{0+\Delta t})}} - \mathbf{x_t^{t_{0+\Delta t}}})$
$\delta\mathbf{e1}$	$\mathbf{2}(\mathbf{x_a^{t_0}} - \mathbf{x_f^{t_0}})^{\mathbf{T}}\mathbf{K}^{\mathbf{T}}\mathbf{M_f^T}\mathbf{C}(\mathbf{x_a^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}})$
$\delta\mathbf{e2}$	$(\mathbf{x_a^{t_0}} - \mathbf{x_f^{t_0}})^{\mathbf{T}}[\mathbf{M_f^T}\mathbf{C}(\mathbf{x_a^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}}) + \mathbf{M_a^T}\mathbf{C}(\mathbf{x_f^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}})]$
$\delta\mathbf{e3}$	$(\mathbf{x_a^{t_0}} + \mathbf{x_f^{t_0}})^{\mathbf{T}}[\mathbf{M_f^T}\mathbf{C}(\mathbf{x_a^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}}) + \mathbf{M_a^T}\mathbf{C}(\mathbf{x_a^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}})]$
$\delta\mathbf{e4}$	$(\mathbf{x_a^{t_0}} - \mathbf{x_f^{t_0}})^{\mathbf{T}}[\mathbf{M_a^T}\mathbf{C}(\mathbf{x_f^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}}) + \mathbf{M_a^T}\mathbf{C}(\mathbf{x_a^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}})]$
$\delta\mathbf{e5}$	$(\mathbf{x_a^{t_0}} - \mathbf{x_f^{t_0}})^{\mathbf{T}}[\mathbf{M_f^T}\mathbf{C}(\mathbf{x_f^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}}) + \mathbf{M_f^T}\mathbf{C}(\mathbf{x_a^{\Delta t+t_0}} - \mathbf{x_t^{\Delta t+t_0}})]$

Here, $\mathbf{M_a} = \frac{\partial x_a^{t_0+\Delta t}}{\partial x_a^{t_0}}$ & $\mathbf{M_f} = \frac{\partial x_f^{t_0+\Delta t}}{\partial x_f^{t_0}}$ are the resolvent matrices of the tangent linear version of the WRF model and $\mathbf{M_a^T}$, $\mathbf{M_f^T}$ are the adjoint WRF matrices, that transfer the sensitivity gradients from forecast time to assimilation time. For example, $\delta\mathbf{e5}$ differs from $\delta\mathbf{e4}$ only by the choice of the adjoint matrix $\mathbf{M_f^T}$ which replaces the adjoint matrix $\mathbf{M_a^T}$. This means that for $\delta\mathbf{e5}$ the linearization is done along the trajectory of the simple forecast $\mathbf{x_f^{\Delta t+t_0}}$ instead of the forecast initiated from the analysis $\mathbf{x_a^{\Delta t+t_0}}$.

The augmented version of the error approximation was adopted for this study,

$$\delta\mathbf{e4} = (\mathbf{x_a^{t_0}} - \mathbf{x_f^{t_0}})^{\mathbf{T}}[\mathbf{M_a^T}\mathbf{C}(\mathbf{x_f^{t_{0+\Delta t}}} - \mathbf{x_t^{t_{0+\Delta t}}}) + \mathbf{M_a^T}\mathbf{C}(\mathbf{x_a^{t_{0+\Delta t}}} - \mathbf{x_t^{t_{0+\Delta t}}})]. \tag{2.20}$$

This version was found by Gelaro et al. [2007] to be a better approximation to the non-linear error $\mathbf{\Delta e}$ compared to the other approximations ($\delta\mathbf{e1}, \delta\mathbf{e2}, \delta\mathbf{e3}, \delta\mathbf{e5}$). This was achieved by exactly calculating the values of the non-linear error $\mathbf{\Delta e}$ and its approximations according to the formulas in Table 2.1. This result was also confirmed in this study, where for some selected days the relative error between $\mathbf{\Delta e}$ and its approximations was exactly calculated for the European domain.

In order to verify the results qualitatively, the traditional but rather computational expensive OSEs experiments were performed for the targeted domain, time span, and observations. The details of how the experiments were conducted and the results can be found in Section 5.2.1.

Chapter 3

Weather and Research Forecasting Model

The WRF model is a state of the art numerical weather prediction system for atmospheric research (Skamarock et al. [2008]). The development of WRF was a collective effort from a number of institutions, namely the National Center for Atmospheric Research (NCAR), the National Centers for Environmental Prediction (NCEP), Forecast Systems Laboratory (FSL), Air Force Weather Agency (AFWA), Federal Aviation Administration (FAA) at the late $90's$. In this work the WRF-ARW model (Advanced Research WRF) is utilized, developed from NCAR's Mesoscale and Microscale Meteorology Laboratory.

The system consists of the dynamical solver, dynamic/numeric options, physics schemes and the data assimilation component. Specifically the ARW solves fully compressible, non-hydrostatic Euler equations (advection, Coriolis, buoyancy etc) and the prognostic variables include velocity components in Cartesian coordinates, perturbation temperature, perturbation geopotential and perturbation surface of dry air. Optionally, TKE (Turbulent Kinetic Energy) and scalars such as cloud water/ice mixing ratio, water vapor mixing ratio and chemical species and traces can be also prognosed. The used grid is the Arakawa-C grid staggering stensil as shown in Fig. 3.1.

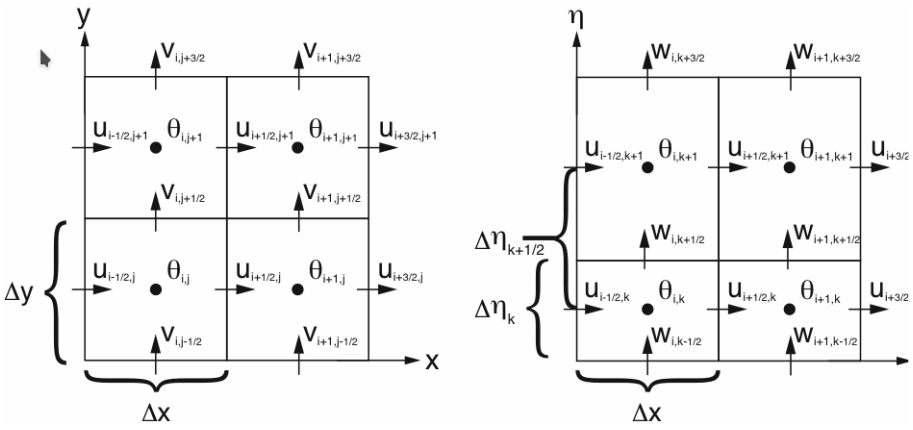


Figure 3.1: On the left, the horizontal and on the right the vertical Arakawa C-grid configuration is depicted. At the center of the left face are the zonal velocities in X direction and at the center of the front face are the meridional velocities in Y direction. At the center of the bottom face are the vertical W points in Z direction and at the center are the Θ points (Skamarock et al. [2008]).

Scalar state variables (thermodynamical variables such as potential temperature and pressure) are calculated at Θ points and the horizontal wind components are stored in X stagger (u) and Y stagger (v) dynamics. The vertical wind component w is stored in the Z stagger dynamics at half levels. As can be seen in Fig. 3.2, the vertical coordinate is not evenly spaced because it is a terrain following, dry hydrostatic pressure with a constant pressure $p_{\eta t}$ at the top of the model.

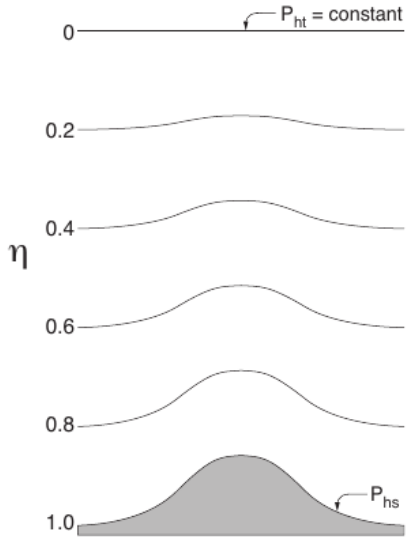


Figure 3.2: The equations in ARW are formulated using the terrain following hydrostatic-pressure defined as $\eta = (p_\eta - p_{\eta t})/(p_{\eta s} - p_{\eta t})$ (Skamarock et al. [2008]). p_η is the hydrostatic pressure at the level of interest and $p_{\eta s}$, $p_{\eta t}$ are the surface and top of model hydrostatic pressure, respectively.

The temporal discretization is achieved by the Runge-Kutta time integration scheme, where a set of ordinary differential equations are integrated in three time

steps using a predictor-corrector formulation (Skamarock et al. [2008]).

For the real-data simulations, the WRF Preprocessing System (WPS) is utilized. Its main tasks include defining the simulation domain, interpolating terrestrial data to the simulation domain and interpolating the meteorological data from another model to the specified simulation grid. Here, the analyses from the Global Forecasting System (GFS) are used to initialize WRF. GFS is operated by NCEP and is run with a 0.5° resolution with an analysis available every, 00, 06, 12 and 18 UTC. The work flow of the preprocessing system is shown in the following Fig. 3.3.

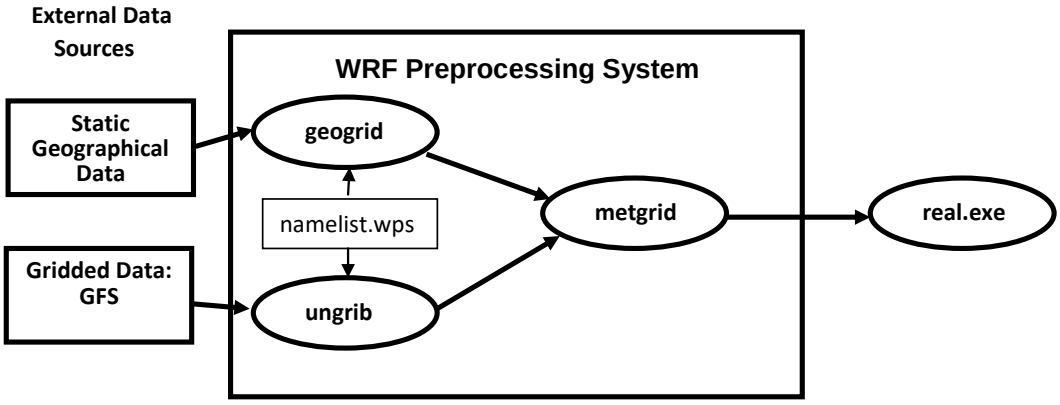


Figure 3.3: WPS work-flow. The interpolated data from another model (e.g. GFS) into the simulation domain are passed to the real.exe program responsible for vertically interpolating the meteorological data to the simulation model levels (Skamarock et al. [2008]).

3.1 Variational data assimilation in WRF model

The WRF Data Assimilation System (WRFDA) with 3DVAR, are utilized in this work. 3DVAR provides the maximum likelihood (minimum variance) estimate of the true atmospheric state, the analysis, given a previous forecast x_b, often termed "background", and observations y. This is accomplished by the iterative minimization of the following predefined cost function:

$$\mathbf{J(x) = J^b + J^o = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(y - Hx)^T R^{-1}(y - Hx)}. \quad (3.1)$$

The 3DVAR procedure is applied in order to provide the analysis for the FSO scheme, which then calculates the contribution of the assimilated observations on the short-term forecasts (see Chapter 2, Fig. 2.1). In Eq. (3.1), $\mathbf{H}$ is the observation operator that interpolates from model space to observation space. The distance of the observations from the first guess model equivalent is called the innovation vector:

$$\mathbf{d = y - Hx_b}, \quad (3.2)$$

where $\mathbf{y}$ is a vector containing the observations one wishes to assimilate and $\mathbf{x_b}$ contains the model quantities. $\mathbf{R}$ and $\mathbf{B}$ are the observation and background error covariance matrices, respectively. The quadratic cost function assumes that both the observation error covariance matrix $\mathbf{R}$ and the background error covariance $\mathbf{B}$ are described using Gaussian probability density functions with zero mean error and correlations between observation and background errors are neglected (Barker et al. [2003]).

Moreover to calculate the background J^b component of the cost function, a preconditioning through a control variable ν transform is taking place. Taking into consideration that the degrees of freedom involved in the calculations for a NWP model is of $n \simeq 10^7$ magnitude, the J^b term of the cost function demands $\simeq O(n^2)$ calculations (Barker et al. [2003]). To reduce the number of calculations, a variable transform is chosen as $x' = U\nu$, where $x' = x - x_b$. U is chosen so that $B = UU^T$ holds true. With this transformation the background error covariance matrix is diagonalized and the number of calculations can be reduced to $O(n)$ (Barker et al. [2003]).
The configuration of the 3DVAR assimilation system in WRF is depicted in Fig. 3.4.

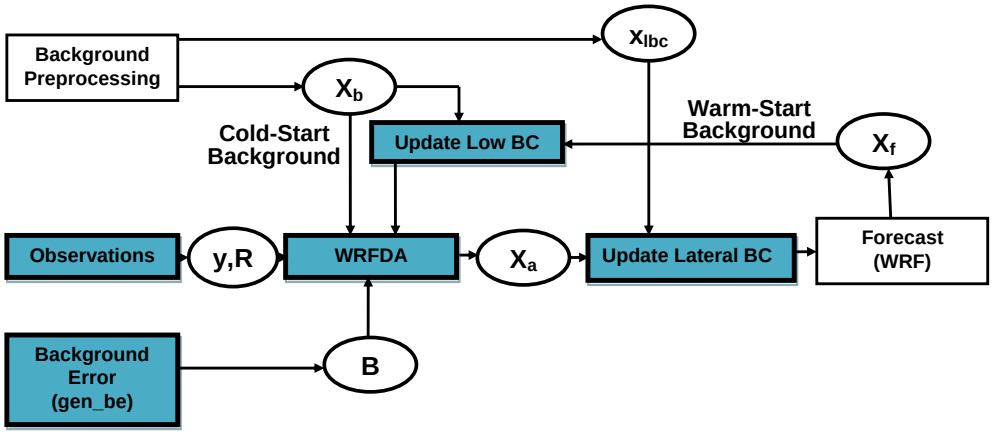


Figure 3.4: Flow-chart of WRFDA configuration where : $\mathbf{x_b}$ is the first model guess, initiated either from WRF forecast or from WPS output. $\mathbf{x_{lbc}}$ are the lateral boundary conditions calculated from WPS. $\mathbf{x_a}$ is the analysis and $\mathbf{x_f}$ is the forecast. $\mathbf{y}$ are the observations, $\mathbf{B}$ is background error covariance matrix and $\mathbf{R}$ is the observation error covariance matrix (Skamarock et al. [2008]).

The domain specific background error covariance matrix, $\mathbf{B}$, is generated for the simulations in this work with the National Meteorological Center (NMC) method, with the help of the program `gen_be` (Fig. 3.4). More specific, in the NMC method (Parrish and Derber [1992]) the background error $\varepsilon_\mathbf{b}$ is approximated by averaged differences between forecasts with the same verification time but different initiation

time (Skamarock et al. [2008]). This means two forecasts are initiated, $x_{(-12,+12)}$ and $x_{(0,+12)}$, with a time difference of 12 hours. Then the background error can be written as,

$$B = \overline{(x_b - x_t)(x_b - x_t)^T} = \overline{\varepsilon_b \varepsilon_b^T} \approx \overline{(x_{(-12,+12)} - x_{(0,+12)})(x_{(-12,+12)} - x_{(0,+12)})^T}. \tag{3.3}$$

The over-bars in Eq. (3.3) denote an average over time. Input data were WRF forecasts and a month long dataset was used to determine the background error covariance tailored to the desired domain (Skamarock et al. [2008]).

The ingestion of observations into WRFDA is shown in Fig. 3.5. A user-defined time slot (1 hour in this study) is used. This means, that observations of ±1 are all assimilated together, after having passed the quality control procedures successfully.

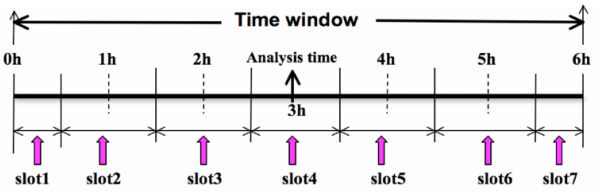


Figure 3.5: Observations are ingested into WRFDA within time slots. If analysis time is at 3h, then a slot of 1 hour is defined and all available observations within that period are assimilated.

To assimilate the satellite radiances (see Section 4.1.1 for information on the examined satellite observations and Section 4.2.1 for radiance assimilation), the Community Radiative Transfer Model (CRTM), which is embedded in WRFDA is utilized. CRTM is developed at the Joint Center for Satellite Data Assimilation (JCSDA) and it is a sensor-based radiative transfer model. It supports more than 100 sensors and has 4 basic modules (JCSDA [2017]):

- Radiative transfer solver
- Surface emission and reflection
- Cloud and aerosol absorption and scattering
- Gaseous transmittance

In the core of the model lies the radiative transfer equation (RTE), which describes the attenuation of the electromagnetic radiation as it propagates through the atmosphere due to absorption, emission and scattering. The radiance observations used in this work are in the thermal IR (Infrared) spectrum from 3.63 to 15.5 μm. The available data for this study are cloud-cleared (see Section 4.1), thus the solution for the

monochromatic intensity $\mathbf{I}(\mu)$ for the non-scattering RTE is (Han [2006]):

$$I(\mu) = [r\int_0^{\tau_N} B(T)dT_d(\tau',\mu_d) + r_\otimes \frac{F_\otimes}{\pi} T_d(0,\mu_\otimes) + \varepsilon B(T_s)]T_u(\tau_N,\mu) - \int_0^{\tau_N} B(T)dT_u(\tau',\mu) \quad (3.4)$$

In Eq. (3.4), $\tau_{\mathbf{N}}$ stands for the optical depth of the atmosphere and $\mu = cos(\theta)$, where θ is the zenith angle. $\mathbf{T_d}(\tau', \mu_\mathbf{d})$ and $\mathbf{T_u}(\tau, \mu)$ are the downwelling and upwelling transmittances, respectively. $\mathbf{F}_\otimes$ is the solar irradiance, $\mathbf{r}$ and $\mathbf{r}_\otimes$ is the surface reflectivity and ε is the surface emissivity. $\mathbf{B}$ i the Planck function. For simplicity, the wavelength subscript is omitted. On the right hand of Eq. (3.4) the four terms correspond to:

1. Downwelling atmospheric radiation, reflected by Earth's surface.
2. Surface reflected solar radiation.
3. Surface emission at skin temperature $\mathbf{T_s}$.
4. Contribution by the atmospheric upwelling radiation.

The solution for each radiance channel can be found by integrating both sides of Eq. (3.4) with the channel's Spectral Response Function (SFR). This is done under the assumption that $\varepsilon, \mathbf{r}, \mathbf{r}_\otimes$ and $\mathbf{B}$ do not fluctuate significantly within the spectral band of the sensor channel. A discrete form of Eq. (3.4) is then obtained. More details can be found in Han [2006]. Moreover, instead of the intensity $\mathbf{I}(\mu)$, it is common to work with brightness temperatures (T_b) in units of Kelvin (K). T_b, refers to the equivalent blackbody temperature of the monochromatic intensity $\mathbf{I}(\mu)$ and can be derived as :

$$T_b = B^{-1}(I(\mu)) \quad (3.5)$$

where $\mathbf{B^{-1}}$ is the inverse of the Planck function and is applied to the observed by the sensor radiance $\mathbf{I}(\mu)$.

To calculate the first guess or model equivalent observations for the 3DVAR cost function (Eq. (3.1)), the forward model of the CRTM calculates the corresponding T_b as seen by a satellite sensor, given as input temperature (T), absorber profiles (q) and surface temperature (T_s) (JCSDA [2017]). Then by minimization of the cost function, the minimum is sought where :

$$\begin{aligned}
&\frac{dJ}{dx} = 0, \\
&\frac{1}{2}B^{-1}2(x - x_b) + \frac{1}{2}R^{-1}2(y - Hx)\frac{d(y - Hx)}{dx} = 0, \\
&B^{-1}(x - x_b) - H^T R^{-1}(y - Hx) = 0, \\
&where, \\
&H^T = \{h_{i,j}\}^T = \{\frac{\partial y_i}{\partial x_j}\}^T
\end{aligned} \tag{3.6}$$

is the transpose of the Jacobian matrix or adjoint operator. $\mathbf{H}$ is a linearisation of the forward model about the state variables $\mathbf{x_j}(\mathbf{T}, \mathbf{q}, \mathbf{Ts}, ..)$. Given as input, perturbations of the state variables it outputs the expected perturbations to the different channels $\mathbf{y_i}(\mathbf{T_B}, \mathbf{I}(\mu))$. $\mathbf{H}^T$ is the transpose of the tangent linear operator $\mathbf{H}$ mapping input perturbations $\mathbf{y_i}$ to output perturbations of state variables $\mathbf{x_j}$.

Chapter 4

Experiment set-up

With priority given to exceptionally poor predictability days of wind and solar power, a two weeks period (01.-15.08.2014) was selected to perform the simulations for this study. This time frame includes the extreme wind error event on 09.08.2014, as identified by Good [2017]. Each day, four assimilation runs by WRFDA were performed for 06, 12, 18 and 00 UTC. This makes a total of 60 assimilation processes performed, a satisfying number to verify the FSO algorithm against the OSEs. To serve the needs of energy-meteorology, 6-hour forecasts (intraday forecasts -up to 6-hour forecast horizon-) were performed. The impact of the assimilated observations on the 6-hour forecasts was then calculated. The experiment was run in a cycling mode, meaning the WRF forecasts initialized from the analysis on time $\mathbf{t}$ were used first for computing the adjoint forcing at verification time $\mathbf{t+6}$ and as background for the assimilation in the next cycle (Fig. 4.1).

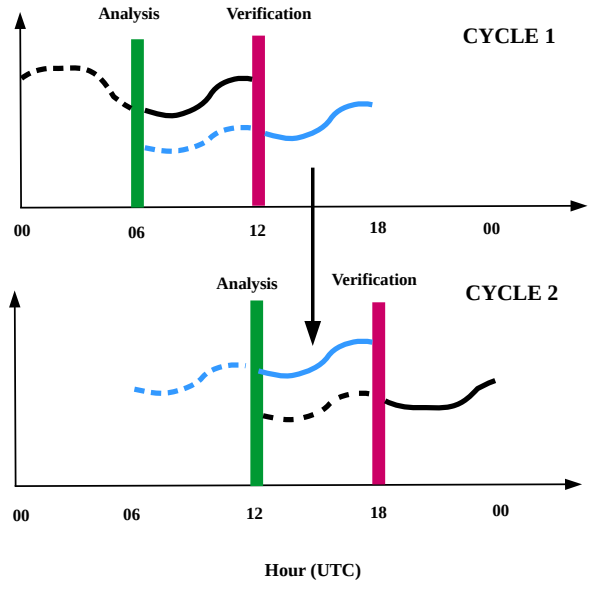


Figure 4.1: The cycling mode of the experiment set-up. The forecast at verification time 12 UTC (CYCLE 1) that was initiated from the analysis (blue dashed lines) becomes the background information for the analysis in CYCLE 2.

The scheme was set-up in the Jülich Research on Exascale Cluster Architectures (JURECA, Jülich Supercomputing Centre [2018]). The different Message Passing Interface (MPI) tasks were managed through a shell script. Within one cycle of observation impact calculation, the data assimilation algorithm was run for the minimization of the 3DVAR cost function and for the calculation of the analysis error covariance matrix (needed for the impact calculation as described in Chapter 2, Eq. (2.10)). Two forward and two adjoint runs followed. In total, 60 impact cycles between 01-15.08.2014 were performed. The configuration of the model can be found in Table 4.1 and the domain is depicted in Fig. 4.2. The parameterizations (physics, boundary layer, etc) for the runs can be found in Appendix A.3, where the complete namelist options for the forward, assimilation and adjoint runs are listed.

Table 4.1: WRF model configuration

Model Version	3.8
Map Projection	Lambert Conformal
Central Point	$54°N, 12.5°W$
Horizontal resolution	15 km
Number of horizontal grid points	250 x 250
Vertical layers	41

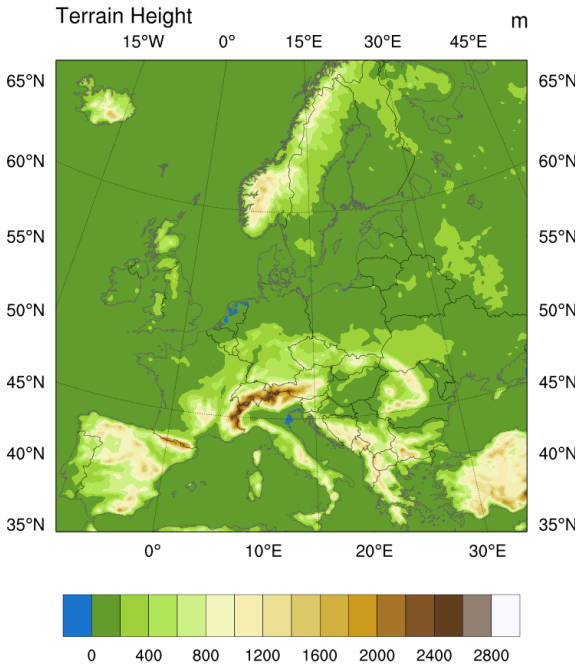


Figure 4.2: WRF domain used in the study. The terrain height is depicted in meters (m).

The computational needs are summarized in the following Table 4.2. The computational demands for the adjoint runs stand out, both in time and cores needed for

their execution. The number of core hours per FSO experiment cycle is the time needed for the execution of all programs of the scheme, multiplied by the used cores. The total core hours consumed is the product of the core hours per FSO cycle with the number of FSO cycles performed in the fortnight.

Table 4.2: Computational resources on JURECA. One FSO cycle demands the execution of WRF forward model, assimilation and adjoint model. The calculation of the impact is done with the help of the WRF assimilation model.

domain	$15km \times 15km$
number of FSO runs (15 days x 4 assimilation runs per day)	60
number of cores/nodes per FSO experiment cycle	768/16
number of cores/nodes per WRF forward runs (two runs for FSO)	192/4
number of cores/nodes per WRF assimilation run	96/2
number of cores/nodes per WRF impact run	96/2
number of cores/nodes per WRF adjoint runs (two runs for FSO)	384/8
number of core-hours per FSO experiment run	2688
percentage of total time needed for the execution of the forward runs	19,05%
percentage of total time needed for the execution of the assimilation run	7,14%
percentage of total time needed for the execution of the impact run	21,43%
percentage of total time needed for the execution of the adjoint runs	52,38%
total number of core-hours needed for the case study period	161,280

4.1 Observation configuration

Setting the focus on a discussion of space borne versus ground based observation configuration, the spatially dense ground based Surface Synoptic (SYNOP) observations (temperature (T), wind speed (U,V), pressure (P) and humidity (Q)- seeFig. 4.11) are evaluated against the profiles of temperature and humidity provided by the measurements of the radiance channels of the Infrared Atmospheric Sounding Interferometer (IASI), onboard MetOp. For convenience, as seen later on, IASI radiance channels are grouped into three parts:

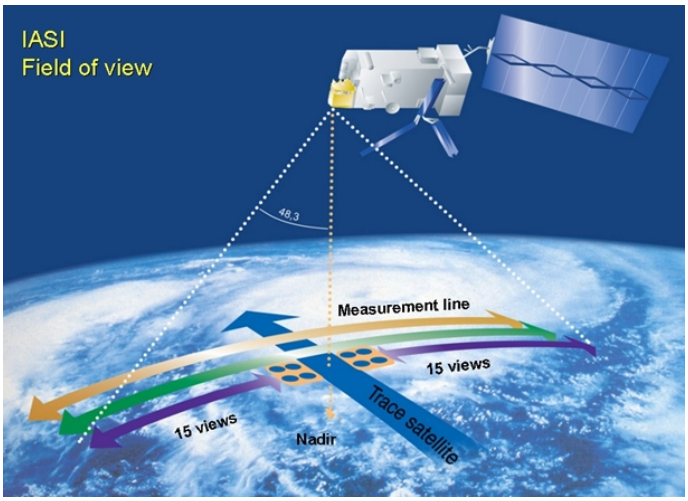
- $IASI\_T$, for channels which are used to derive temperature at a level of the atmosphere,
- $IASI\_W$, for channels positioned at the window region of the infrared spectrum and give information like surface temperature,
- $IASI\_Q$ for channels in the infrared from which humidity is derived.

A description of the IASI and SYNOP data follows in section Section 4.1.1 and Section 4.1.2. Details on the observations that were assimilated in order to provide the reference state (alongside IASI and SYNOP) for FSO are then given in Section 4.2. Focus is then placed on the assimilation of IASI radiance channels and SYNOP in Section 4.2.1 and Section 4.2.2, respectively.

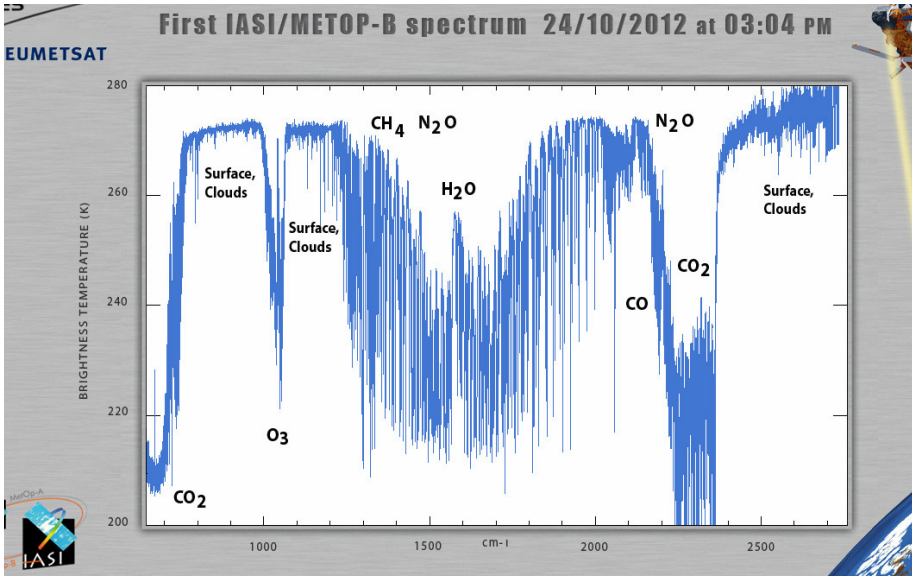
4.1.1 Infrared Atmospheric Sounding Interferometer

The IASI sensor is onboard the polar orbiting MetOp satellites which are operated by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). MetOp is presently operating a series of three polar orbiting meteorological satellites that form the EUMETSAT Polar System (EPS). It carries a set of 'heritage' instruments provided by the United States and a new generation of European instruments that offer improved remote sensing capabilities to both meteorologists and climatologists (Metop [2017]). Metop-A (launched on 2006) and Metop-B (launched on 2012) are in a low orbit altitude of 817 kilometers, providing observations of the global atmosphere, continents and oceans. They are sun-synchronous, mid-morning orbiting (9:30 Local Solar Time equator crossing, descending node) and provide global observations twice a day.

IASI consists of an across-track scanning system of a symmetrically scan range of $\pm 48°20'$. A scan line covers 30 scan positions towards the Earth and 2 calibration views, the one pointing into deep space and the other on an internal black body (EUMETSAT [2017]). The elementary (or effective) field of view (EFOV) is the useful field of view at each scan position. Each EFOV consists of a 2**x**2 matrix of the so-called instantaneous fields of view (IFOV). Each IFOV corresponds to a ground resolution of 12 km at nadir and a satellite altitude of 819 km (EUMETSAT [2017]). The swath width is approximately 2.2000 km (Fig. 4.3(a)).



(a)



(b)

Figure 4.3: (a): IASI on-board MetOp. The trace of the satellite and the Field of View (FOV) is depicted. (b) : First IASI/METOP-B spectrum on 24.10.2012 at 03:04 pm
Reference:
(a)://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/IASI/index.html
(b)://www.eumetsat.int/website/home/Data/RegionalDataServiceEARS/EARSIASI/index.html?lang=EN

The main objective of IASI is to provide atmospheric emission spectra of the infrared part of the electromagnetic spectrum, from which the temperature and humidity profiles can be derived in high vertical resolution and accuracy. In order to retrieve the atmospheric state that produces the measured spectrum, the forward model (CRTM in this study) that describes the radiative transfer through the atmosphere is inverted. The atmospheric components and surface information retrieved from IASI measurements can be seen in Fig. 4.3b and are listed in Table 4.3 (Chalon et al. [2001]).

The measured spectral range extends from 645 to 2740 cm^{-1} (or 15.5 to 3.62 μm) with sampling every 0.25 cm^{-1} that leads to a total of 8461 radiance channels.

Table 4.3: Spectral range given in units of cm^{-1} and μm along with the corresponding IASI application (Chalon et al. [2001]).

Spectral range	Absorption band	IASI application
650 to 770 cm^{-1} or 15.4 to 12.9 μm	CO_2	Temperature
790 to 980 cm^{-1} or 12.6 to 10.2 μm	Atmospheric Window	Surface and Cloud properties
1000 to 1070 cm^{-1} or 10.0 to 9.3 μm	O_3	O_3 sounding
1080 to 1150 cm^{-1} or 9.2 to 8.7 μm	Atmospheric Window	Surface and Cloud properties
1210 to 1650 cm^{-1} or 8.3 to 6.0 μm	H_2O	Humidity profile and N_2O column amount
2100 to 2150 cm^{-1} or 4.8 to 4.7 μm	CO	CO column amount
2150 to 2250 cm^{-1} or 4.7 to 4.4 μm	N_2O and CO_2	Temperature profile, N_2O column amount
2350 to 2420 cm^{-1} or 4.2 to 4.1 μm	CO_2	Temperature profile
2420 to 2700 cm^{-1} or 4.1 to 3.7 μm	Atmospheric Window	Surface and cloud properties
2700 to 2760 cm^{-1} or 3.7 to 3.6 μm	CH_4	CH_4 column amount

This range includes the CO_2 absorption band at $15\mu m$ from which the temperature profiles are derived, the strong water vapor ν_2 (bending mode of water molecule) absorption band and the $9.6\mu m$ Ozone band (Clerbaux et al. [2009]). Generally, the derivation of geophysical parameters of the L2 products, such as trace gases concentrations, temperature and water vapor profiles from radiance measurements is performed in collaboration with the Advanced Very High Resolution Radiometer (AVHRR), ATOVS and data from NWP's (Metop [2017]).

The IASI data used for this study are from NCAR's Research Data Archive (data set ds099.0, NCEP [2015]). Through e-mail communication with the curator of this data base, it was found that the satellite data assimilated by the Global Data Assimilation System (GDAS) had been archived. Thus, the data had been quality-controlled and thinned by NCEP for their data assimilation system. The provided data set consists of a subset of 616 channels out of the 8461 available channels, originally distributed by the National Environmental Satellite, Data, and Information Service (NESDIS) (Hilton et al. [2012]). The distributed data are thinned by sub-sampling FOVs and channels using the same method developed

for the Atmospheric Infrared Sounder (AIRS) in Goldberg et al. [2003]. The data are cloud cleared according to Eyre and Menzel [1989], which results in less than 120 IFOVs per scan line in case of cloud presence. To do so, according to their methodology, they seek to find the pressure and cloud fraction, for which the deviation of the calculated cloud from model (over a number of channels) and the observation is minimized. It is assumed that the cloud cover has only a single layer, with unit emissivity within the FOV of the instrument. The affected IFOVs are assigned with a cloud flag and are dismissed by WRFDA. The reduced number of radiance channels were selected based on the spectral properties of the channels, by preserving those with the highest spectral purity and those whose sensitivity function structure provide the best vertical coverage (Gambacorta and Barnet [2013]). For example, channels with a sharper sensitivity function were preferred over broadly structured ones, as they provide better vertical resolution.

Operationally, GDAS (operated by NCEP) assimilates and monitors 167, out of the 616 radiance channels. The majority of these radiance channels lies around the long wave $\mathbf{CO_2}$ band (near $15\mu m$), providing information on temperature. These 167 channels are also a subset of the 300 channels proposed by Collard [2007] and used by a number of NWP centers (Hilton et al. [2012]), with the exception of 2 temperature channels (index: 300 and 326) and in the ozone band (index:1666). These 167 channels, along with the provided observation errors (to be used in the observation covariance matrix $\mathbf{R}$-see Eq. (3.1)) were the initial basis of IASI assimilation with WRF in this study. As a result of WRF not predicting O_3 but only using it as climatology, channels situated around this absorption band were eventually omitted in the assimilation. Channels in the range of 2000 $cm^{-}1$ - 2760 cm^{-1} or 3.62 μm-5.00 μm are not selected because of high instrument noise and influence by solar reflection during the day, making them difficult to process (Hilton et al. [2012]). The original 167 channels provided by NCEP are depicted in Fig. 4.4 along with the 158 channels finally used in this study. In Fig. 4.5 the observation errors of the 158 channels used for the assimilation procedure is shown. As can be noticed in Fig. 4.4, the same channels as NCEP are used for temperature (102 $IASI\_T$ channels) and surface (47 $IASI\_W$ channels) information. Moreover it is deduced from Fig. 4.5, where the assigned by NCEP observation errors of the radiance channels are plotted, that temperature radiance channels have lower observation errors compared to the other two categories.

In addition to the $IASI\_T$ and $IASI\_W$ channels, nine water vapor channels ($IASI\_Q$, $\approx$ 5 to 8 μm, with the water vapor band centered at 6.3 μm) were added in the assimilation in contrast to the NCEP list of channels (no water vapor channels). Although NCEP has researched the impact of water vapor channels (Jung et al. [2009]), the operationally assimilated channels do not include water vapor. The absence of the water vapor channels in NCEP is due to the difficulties that accompany the assimilation, such as the non-linearity of their Jacobians, their multivariance (solve for both T and Q) and their non-Gaussian distribution (Jung et al. [2009]). The nine water vapor channels (see Fig. 4.6), included for the experiments of this thesis, are selected by Météo-France (Guidard et al. [2011]) and they have been found to improve the French global model and their limited area model on the forecast

range longer than T+60 over Europe. Although more water vapor channels are assimilated and investigated by various NWP centers, the specific ones where selected for testing with the FSO algorithm as their exact wavenumber was available. It is within the scope of this research to use the FSO algorithm and quantitatively assess their impact. The emission weighting functions of the nine water vapor channels, which describe the contribution of the distinct atmospheric levels to the emitted radiation observed from space (Petty [2004]) are plotted in Fig. 4.6.

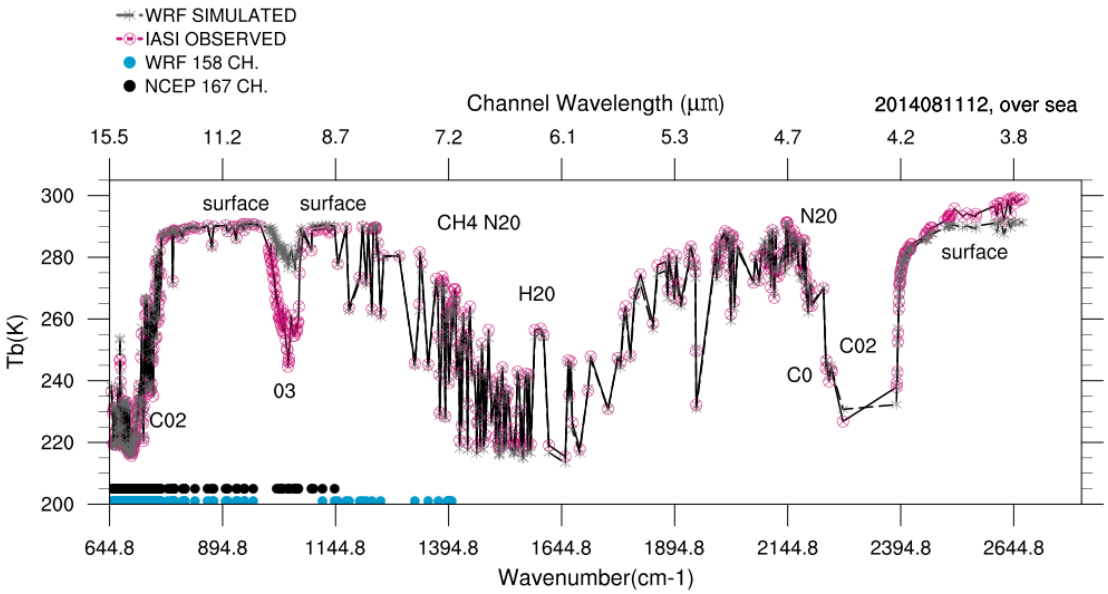


Figure 4.4: The 616 radiance channels as measured by IASI (dark pink crossed circle) and simulated by WRF's CRTM (dark gray asterisk). Black dots represent the position of the 167 channels, assimilated by NCEP and light blue dots are the 158 channels used in this study. The pixel depicted here is located over the Atlantic Ocean.

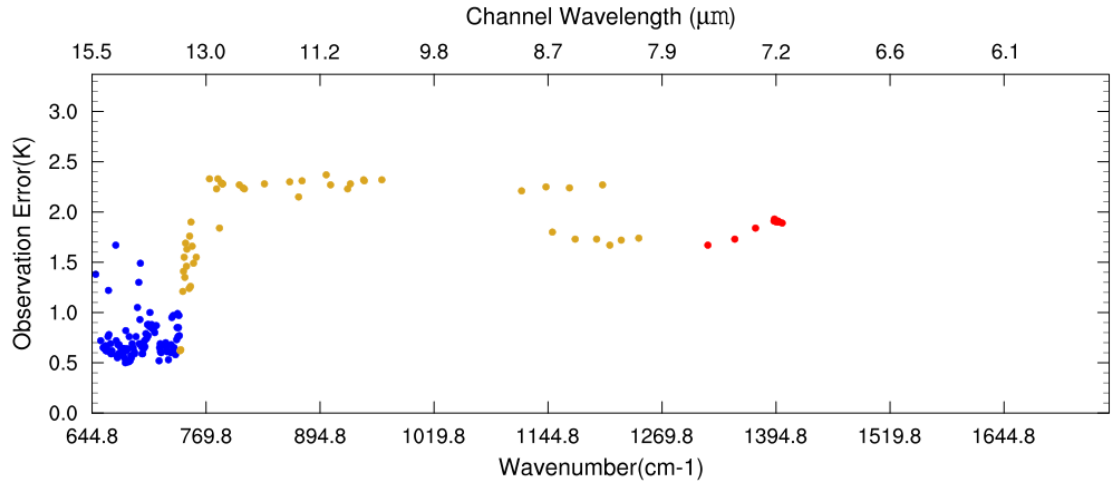


Figure 4.5: Observation errors in Kelvin, as used in the observation covariance matrix **R**, of the assimilated channels. The observation errors of the 102 $IASI\_T$ are shown in blue colors, the 47 $IASI\_W$ in yellow and the 9 $IASI\_Q$ in red.

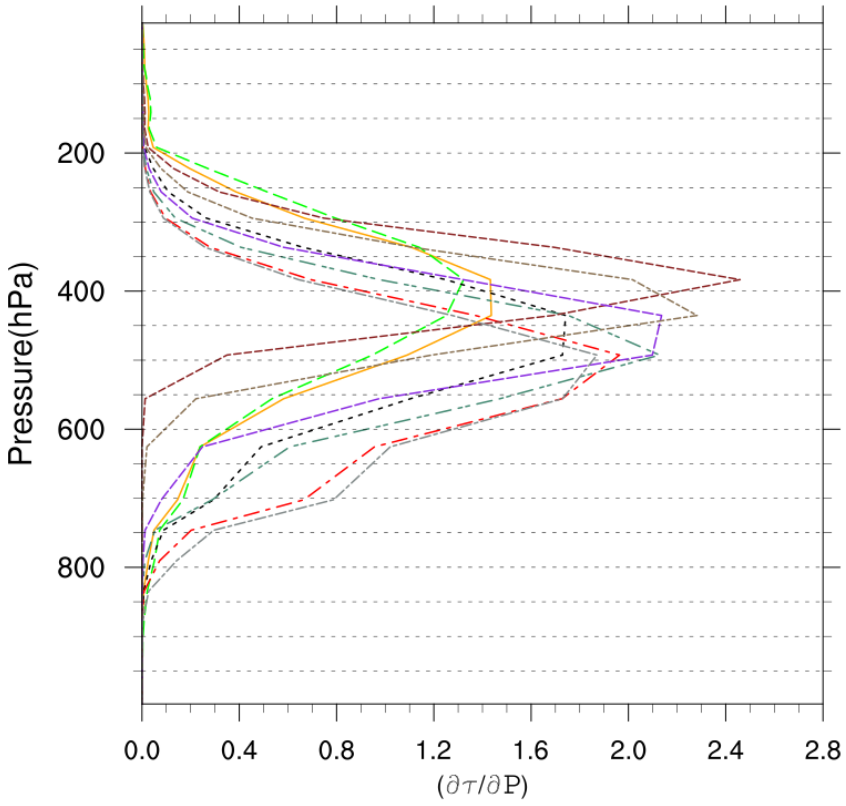
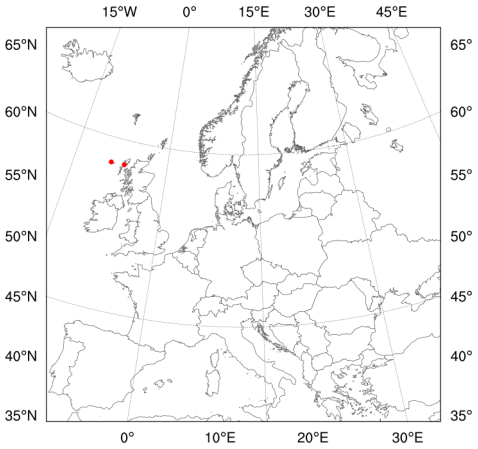
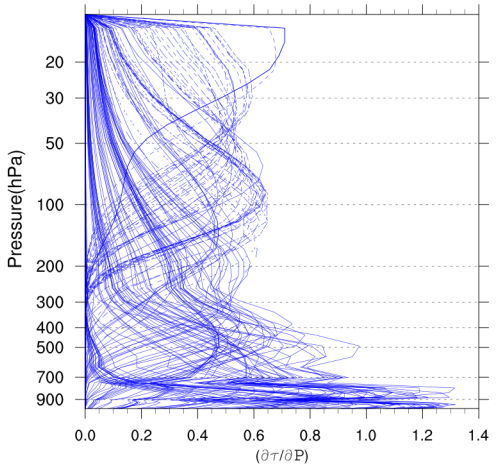


Figure 4.6: The emission weighting functions of the 9 water vapor channels introduced in the assimilation is plotted, as calculated for a pixel over sea by CRTM on 11.08.2014 at 00 UTC.

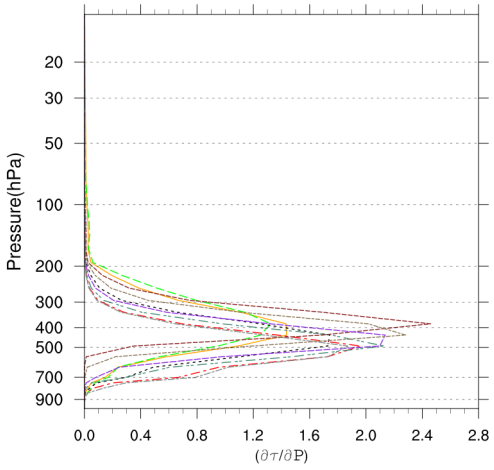
To understand the vertical coverage of all the selected channels, their emission weighting functions are plotted in Fig. 4.7. More specific, the emission weighting functions for the case of a pixel over sea and a pixel over land are plotted. Two plots are generated for each case, one for the temperature channels between ≈ 10 and 15 μm and the other for the 9 water vapor channels defined above. The simulation was done by CRTM for the date 11.08.2014 at 00 UTC. The dense vertical coverage of the troposphere is obvious in Fig. 4.7b, d. From Fig. 4.7e, it is noticed how the water vapor channels are peaking in the lower troposphere, in contrast to the same channels over the sea (Fig. 4.7c). This is an indicator of a dry atmosphere in this case, as water vapor is not well mixed and highly variable. This means that the weighting functions will also vary, peaking in low altitudes in a very dry atmosphere and in the upper troposphere when high altitude clouds are present, or in a humid atmosphere (Petty [2004]).



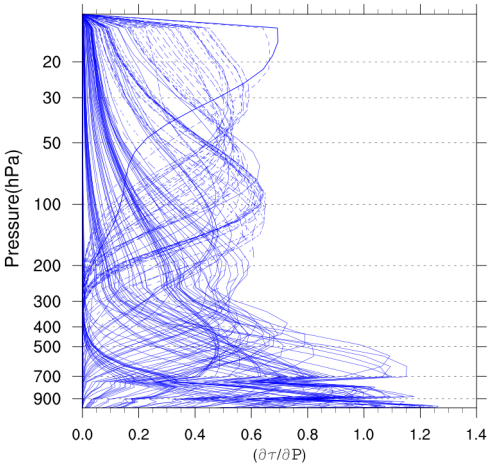
(a) Location of pixels over sea and land



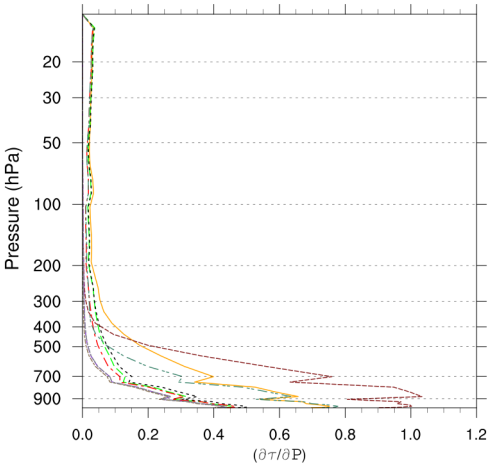
(b) Sea, Temp. channels



(c) Sea, Water vapor channels



(d) Land, Temp. channels



(e) Land, Water vapor channels

Figure 4.7: The emission weighting functions of the temperature and water vapor channels as simulated on 11.08.2014 at 00 UTC by CRTM. In panels (b) and (c), the emission weighting functions of a pixel above sea are plotted and on plots (e) and (f) for a pixel over land. The location of the pixels is shown in plot (a).

4.1.2 Surface Synoptic Observations

SYNOPs are weather reports, made by manned and automated weather stations. They are available in the GTS via the national meteorological organizations in the WMO-based format BUFR. NCEP receives the GTS data and encodes them in a binary format called PREPBUFR that has quality controlled and error assigned observations (Dennis [2018]). In the archived data set, used in this study, observations are grouped around the four synoptic hours (06, 12, 18, 00 UTC) with a time span of ± 3 hours. In the assimilation performed by WRFDA, a time slot of an hour centered around the synoptic times is defined.
There are approximately 2000 SYNOP stations covering Europe. In Fig. 4.8, the station locations that measured temperature (T), humidity (Q), wind (U & V) and pressure (P) at 06 UTC for the period 01.-15.08.2014 are plotted. The number of data from stations varies over the day, indicating the fluctuation of measuring time among the stations but also the possible quality control failure as recognized by NCEP. The maximum number of reporing stations, was found on 06 UTC (2026 stations) and the minimum at 00 UTC (1890 stations).

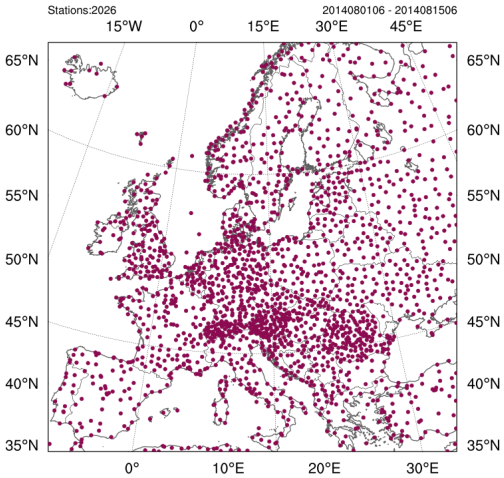


Figure 4.8: Station locations (in total 2026) providing measurements at 06 UTC for the period 01.-15.08.2014.

4.2 Assimilation of observations

Looking back at Chapter 2 and Fig. 2.1, it is clear that a reference state must be provided for the conduction of the experiments. As explained in Chapter 2, the reference state is typically an analysis provided by the same model used to conduct FSO. This analysis is the outcome of the 3DVAR procedure, in which different types of observations are included (along with SYNOP and IASI). Although WRFDA can assimilate a wide range of observation types, a compromise was made to keep the computation limits, resulting in the observation configuration in Table 4.4. FSO calculates then the contribution of the assimilated observations on the 6-hour forecast error. Here, the evaluation of IASI and SYNOP is of interest and FSO is examined on those two observation types. It must be understood, that one assimilation takes place in each FSO cycle and includes all the observations. The analysis serves as reference state in one cycle and on the following cycle it is evaluated by the FSO scheme (see Fig. 4.1). This is considered an advantage of the algorithm, as observations do not need to be removed from the configuration in order to calculate the observation impact of different types on the short-term forecast error.

Table 4.4: Observation types used in the assimilation to produce the reference (true state) analysis. T, RH, P, U and V stand for temperature, relative humidity, pressure and the U, V wind components, respectively.

Data Type	Description	Variables
SYNOP	Surface observations from land and ship stations	U, V, T, P, RH
TEMP	Radiosondes from land and ship	U, V, T, P, RH
METAR	Meteorological Aerodrome Report	U, V, T, P, RH
SHIPS	Measurements onboard ships	U, V, T, P, RH
PILOT	Sondes and wind profiler	U, V
BUOY	Ocean Data Acquisition Systems	U, V, T, P, RH
QSCAT	Scatterometer oceanic surface winds	U, V
IASI	Infrared Atmosperic Sounding Interferometer	radiance assimilation related to T, RH

All observations, with the exception of IASI data, are provided by the 'NCEP ADP Global Upper Air and Surface Weather Observations' and are publicly available (NCEP [2008]). The data set includes land surface, marine surface, radiosonde, and aircraft reports from the Global Telecommunications System (GTS), profiler and US radar derived winds, the Special Sensor Microwave Image (SSM/I) oceanic winds, Total Cloud Water (TCW) retrievals and satellite wind data from the NESDIS. Each data type can be set on/off via the namelist input file of WRFDA. The format of the data is the Prepared Binary Universal Form for the Representation of meteorological data (PREPBUFR), a binary format file that passes through quality control at NCEP making the observations suitable for assimilation.

The data set ds099.0, from NCAR's Research Data Archive (NCEP [2015]) provides the IASI radiance observations (Level 1). A detailed description of the IASI data and how they are processed by NCEP before they are available for public use is

given in Section 4.1.1.

In Fig. 4.9, the different types of assimilated observations, at each assimilation hour (00, 06, 12 and 18 UTC) are plotted as percentage (%) of the total number of assimilated observations for each hour during the case study period 01.-15.08.2014 and the European domain. This means that the bars in Fig. 4.9 at each hour add up to 100%. The observations of interest, SYNOP and IASI have fluctuations of assimilated data over time, with maximum IASI data (compared to other observation types) assimilated at 18 UTC and maximum SYNOP data assimilated at 06 UTC. The absolute numbers of assimilated IASI and SYNOP data for the case study period can be found in Section 5.2, Fig. 5.4.

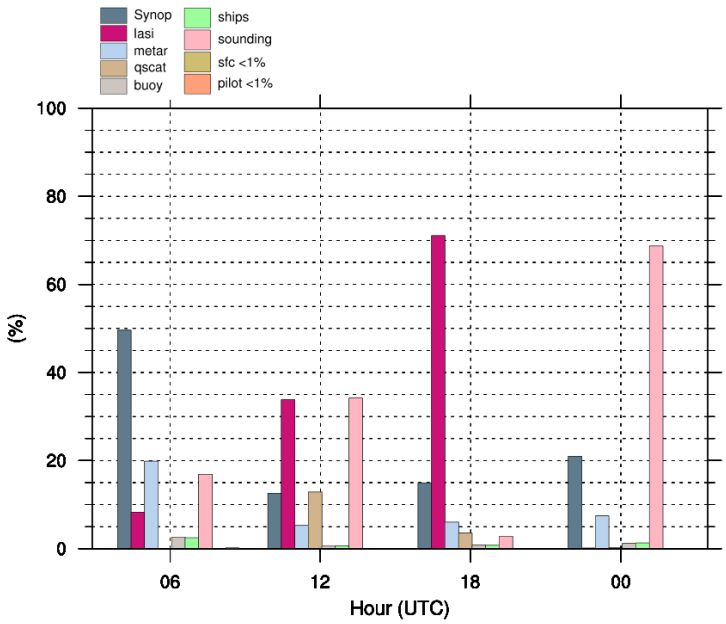
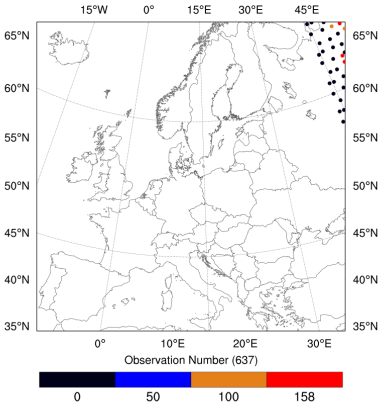


Figure 4.9: Percentage of the assimilated observation types, at each assimilation hour, for the period 01.-15.08.2014 is plotted. The "sfc" observations compose a subcategory of sounding data, that WRFDA defines to include data close to the surface (i.e 10m wind speed, 2m temperature)

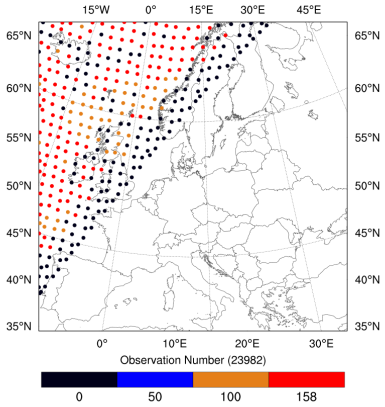
Concerning the spatial coverage of IASI data, the spatial fluctuation is a geometrical side-effect when performing simulations with a limited area model. For example, in Fig. 4.10 the spatial fluctuation at 00, 06, 12 and 18 UTC on 11.08.2014 is demonstrated. The label bars refer to the amount of observations (radiance channels) that are assimilated per pixel. A maximum of 158 wavelength channels can be assimilated per pixel and if this is not the case, it means the observation failed to pass the quality control within WRFDA. A detailed description of the quality control is given in Section 4.2.1.

For the same date as for IASI, the SYNOP observation spatial coverage is plotted in Fig. 4.11. The number of data from stations varies with time, as some stations may not provide regularly data or due to quality control failures. Five SYNOP variables (T, U, V, P, Q) per station can be assimilated. The consistently not assimilated area

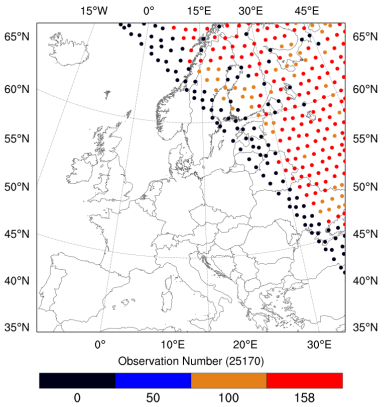
over the Alps (Fig. 4.11) can be explained by the difficulty of WRF -or any limited area model- to model accurately the terrain and the circulation in the boundary layer. For that specific date, clearly either all variables were assimilated or three of them. Out of the five measured variables, mostly the two wind components were rejected. The quality control procedure of WRFDA for SYNOP is described in Section 4.2.2.



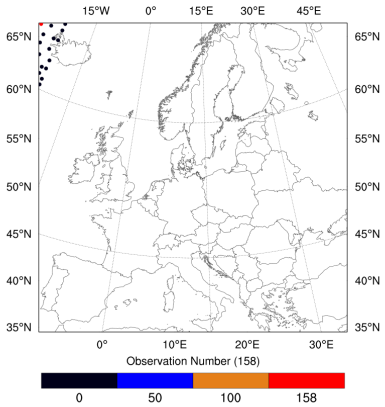
(a) IASI coverage at 06 UTC



(b) IASI coverage at 12 UTC

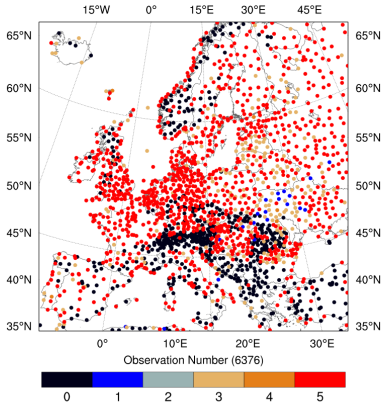


(c) IASI coverage at 18 UTC

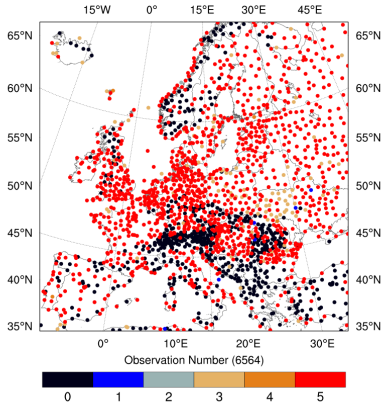


(d) IASI coverage at 00 UTC

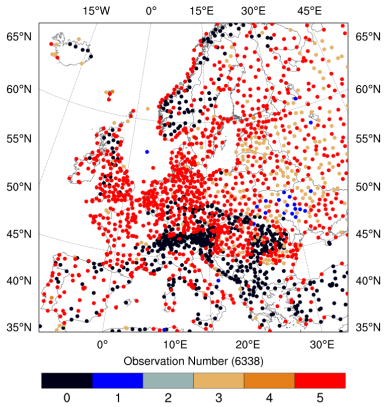
Figure 4.10: IASI spatial coverage at the four assimilation hours, for 11.08.2014. The colored label bar ranges from 0 to 158. This means at each pixel on the map, maximal 158 radiance channels can be assimilated. The data are plotted on the Lambert Conformal projection (not all IFOVs are selected by NCEP (see Section 4.1.1), each dot representing a 12 km horizontal resolution.



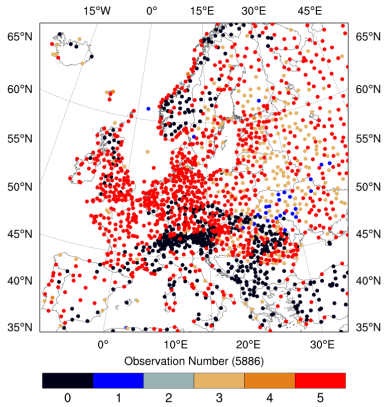
(a) SYNOP coverage at 06 UTC



(b) SYNOP coverage at 12 UTC



(c) SYNOP coverage at 18 UTC



(d) SYNOP coverage at 00 UTC

Figure 4.11: SYNOP spatial coverage at the four assimilation hours, for 11.08.2014. The colored label bar ranges from 0 to 5 indicating the number of assimilated variables: temperature, the two wind components, pressure and specific humidity (converted from measured relative humidity). The range refers to the number of variables that passed the quality control and were assimilated.

4.2.1 IASI assimilation

The preference to directly assimilate T_b rather than profiles of water vapor and temperature lies in the fact that no control of consistency of a priori data is given for the retrieval. Consequently, the observation error covariance of the retrieval is difficult to determine in contrast to the radiance error covariance (Kalnay [2003]). This is also underlined by the choice of most of the European NWPs Centers to assimilate brightness temperature rather than the retrieved profiles (Hilton et al. [2012]), which also provides model consistency as the same NWP will be used for the assimilation.

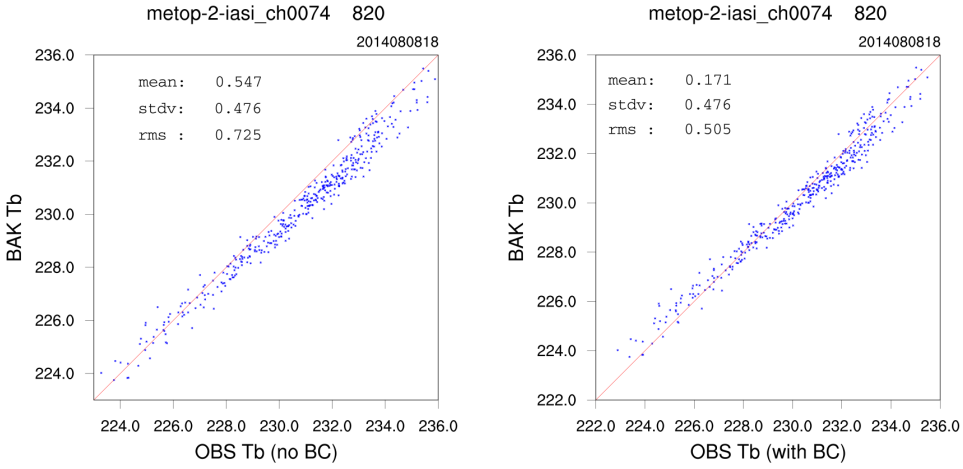
Radiance measurements can be prone to biases. These systematic errors between modeled and actual data can be due to non-modeled effects in the CRTM and the satellite instrument itself (e.g poor calibration) (Auligné et al. [2007]). NWP centers use their assimilation system to correct for these biases, which are not fixed but vary with geographical location, time and scan position(Auligné et al. [2007]). In WRFDA, the Variational Bias Correction (VarBC) module is used to apply bias correction (McNally and Watts [2003]) before the assimilation takes place. The modified observation operator (Liu et al. [2012])

$$\tilde{H}(x,\beta) = H(x) + \beta_0 + \sum_{i=1}^{I_p} \beta_i p_i(x) \tag{4.1}$$

is used. The model simulated brightness temperature $H(x)$ is corrected based on β_0 (constant part of total bias) and the potentially state-dependent predictors p_i and their coefficients β_i which are estimated within a variational minimization process by inclusion in the control variables. The seven parameters included in the predictors p_i are:

1) scan position
2) square of scan position
3) cube of scan position
4) $1.000 - 300$ hPa layer thicknesses
5) $200 - 50$ hPa layer thicknesses
6) surface skin temperature
7) total column water vapor

The effect of this bias correction can be observed in Fig. 4.12, where the scatter plot of the model simulated temperature brightness T_b (BAK) versus the observed T_b (OBS) is plotted. The shift of the distribution around the diagonal is notable in Fig. 4.12b, where the bias correction is applied, along with the reduced root mean square error. The VarBC module is used in every cycle of the FSO algorithm (for the four assimilation runs per day) for the examined fortnight and the European domain.



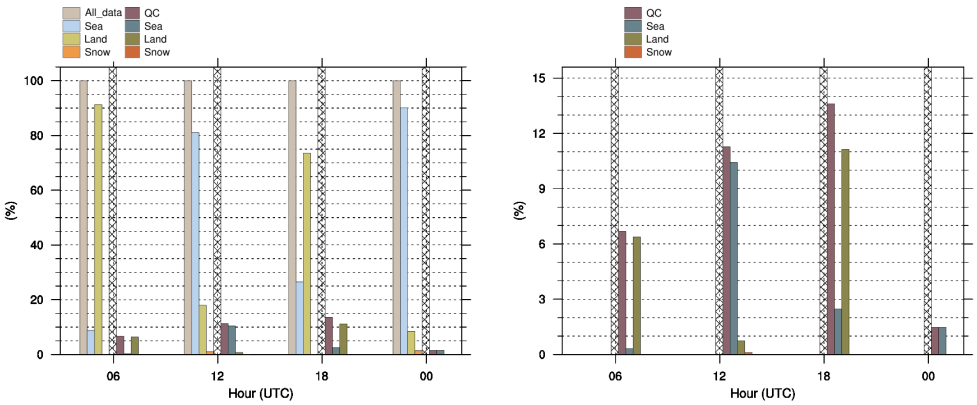
(a) **(b)**

Figure 4.12: Scatter plots of simulated T_b (BAK) and observed T_b (OBS) with (a) no bias correction and (b) with bias correction. The simulation took place on 08.08.2014 at 18 UTC and the 820 observations of the channel 74 of IASI are plotted.

Part of the quality control in WRFDA is then to reject any observation, if the observation minus the bias-corrected radiance exceeds either 15 K or $3\sigma_0$. The σ_0 stands for the brightness temperature specified observation error standard deviation. Other quality control checks, include dismissing channels over mixed surface type, limb observations ($48°20' \leq$ scan positions $\leq 5°$) and channels of wavenumber greater then 2400 cm^{-1}. Even though the data are cloud cleared, an extra elimination of channels is done if the cloud water liquid path is greater than 0.2 kg/m^2, as computed for each layer by the first guess run.

Fig. 4.13a, shows the percentage (%) of IASI data before (left of the crossed column bar) and after (right of the crossed column bar) quality control at each assimilation hour for the period 01.-15.08.2014. There is a distinction of the data in terms of the surface types including sea, land and snow. The total amount of data assimilated (purple column) makes less than 20% of the available data. It is also noticeable that at 12 UTC the majority of the assimilated data is over sea and at 18 UTC over land. Fig. 4.13b is a zoom in the quality controlled assimilated data. There is an increasing percentage of assimilated IASI information from 06 to 18 UTC, with a minimum at 00 UTC. A small percentage ($\leq$ 2 %) of snow surface

type is also assimilated at 12 UTC.



(a) **(b)**

Figure 4.13: (a) The percentage of IASI observations in the period 01.-15.08.2014, before (pale colors, left of the crossed column bar) and after (dark colors, right of the crossed column bar) quality control. With dark purple color are the amount of observations assimilated out of the total number of available observations. A distinction according to surface type (sea, land, snow) is also made. (b) A zoom in plot on the quality controlled observations per assimilation hour.

4.2.2 SYNOP assimilation

During the quality control within WRFDA, a location and time check is performed along with a conversion of the geographical grid to model grid, and a conversion from relative to specific humidity, accompanied by some physical constrains (i.e. negative values of humidity, irrational wind directions). In Fig. 4.14, the total SYNOP data of the two weeks assessment is plotted per synoptic hour. On the right side is the amount of data actually assimilated (after quality control). Approximately 60% of the data is assimilated with an almost equal share between the five SYNOP variables.

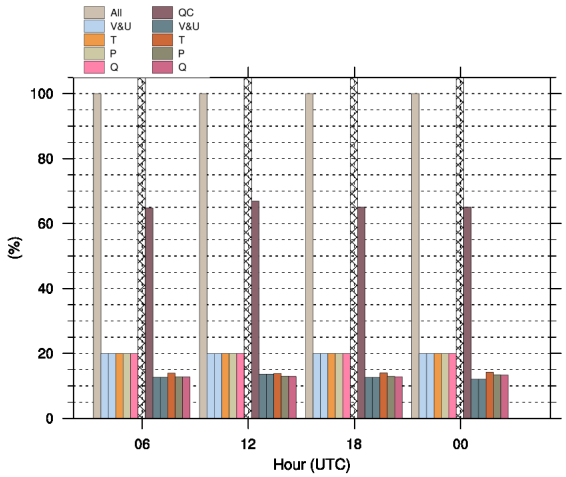
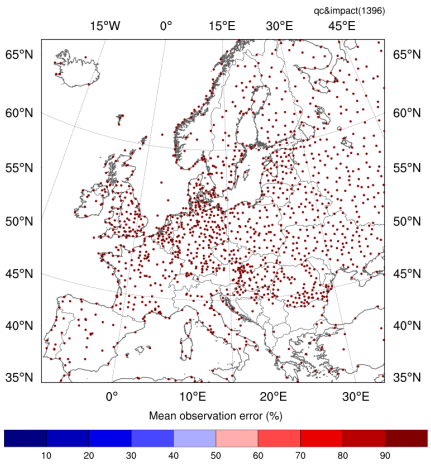
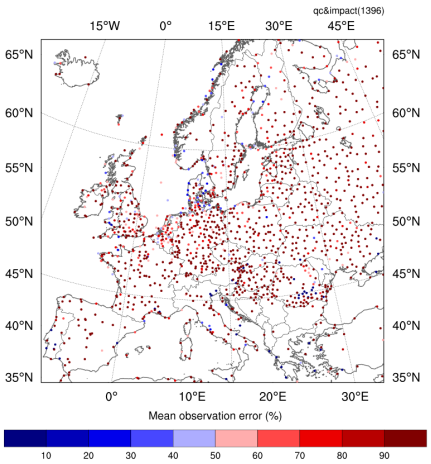


Figure 4.14: Amount of SYNOP data before (pale colors, on the left side of the crossed column bar) and after (dark colors, on the right side of the crossed column bar) quality control.

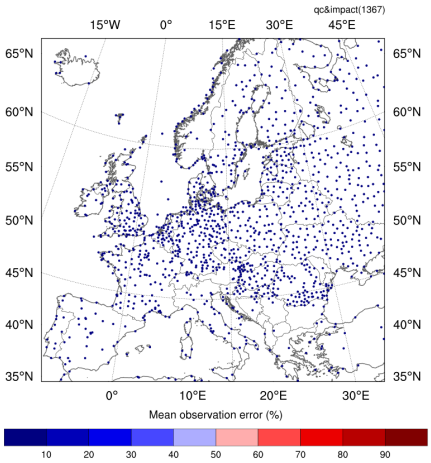
The observation errors of the SYNOP data are assigned by NCEP and distributed along with the measurements. The mean observation error, of the assimilated variables from SYNOP stations is 1.1 m/s for the two wind components, 2 K for temperature, 100 Pa for pressure and 0.0026 kg/kg for specific humidity. To check how representative the assigned errors are, the mean assigned error per variable and per station is calculated, as a percentage of the actual observation values. In Fig. 4.15, the mean observation error (%) is plotted for the SYNOP variables, assimilated at 06 UTC in the examined two weeks period. On average, the temperature and pressure errors (Fig. 4.15c,d) are in the range of 10% of the measurements. Specific humidity observation errors are in between 20-30 % of the actual measurements. The mean observation error for the wind components of 1.1 m/s, attributes a mean percentage of more then 80%. This means the SYNOP measurements were mostly well below 1.1 m/s on average at 06 UTC. A similar qualitative picture is given also for the other analysis hours (12, 18 and 00 UTC).



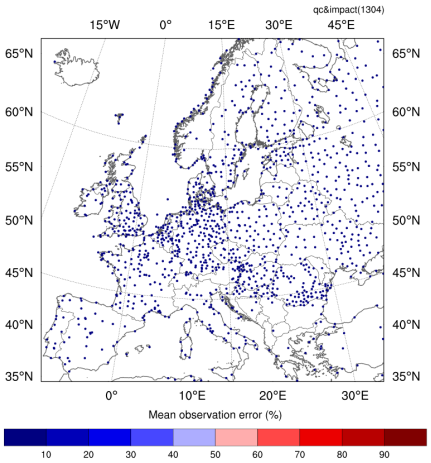
(a) U mean error (%)



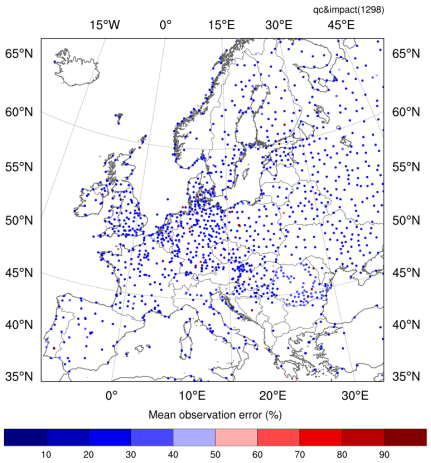
(b) V mean error (%)



(c) T mean error (%)



(d) P mean error (%)



(e) Q mean error (%)

Figure 4.15: The mean observation errors expressed as percentage (%) of the actual observations of (a) the U wind component, (b) the V wind component, (c) the temperature T, (d) the pressure P, (e) the specific humidity Q. The average was calculated for the observations assimilated at 06 UTC for the period 01.-15.08.2014.

Chapter 5

Observation impact results

In this chapter the aggregated FSO results over the defined two weeks period are analyzed. First, the 6-hour non-linear forecast errors and their linear approximations as described in Chapter 2 are examined in Section 5.1 as they indicate the accuracy and performance of the algorithm. The evaluation, for the fortnight, of IASI and SYNOP impact on the 6-hour forecast errors is presented and analyzed in Section 5.2. The identification of cases of poor predictability are then presented in Section 5.3 and evaluated in Section 5.4.

5.1 Non-linear and linear forecast errors

The non-linear 6-hour forecast errors (Eq. (2.14)) and their linearized approximations (Eq. (2.15)) were calculated for the European model domain and for the fortnight 01.-15.08.2014. For the non-linear error, first the differences of the two forecasts from the reference analysis weighted by the energy norm were calculated (gradients at the forecast time). Then the inner product of the gradients with the analysis increment (at forecast time) followed, resulting in one number for the European domain at each assimilation run. The same procedure was followed for the linearized version, with the difference that the gradients were calculated by the runs of the adjoint model from forecast time back to assimilation time. The time series for the two weeks period are presented in Fig. 5.1, where larger absolute negative values indicate larger forecast error reduction. Although the augmented version of the error approximation is adopted (see Chapter 2), Eq. (2.20)), an average of $\approx 60\%$ under-estimation of the non-linear error reduction is found in contrast to other studies that used the same error norm as Gelaro and Zhu [2009] who found an underestimation of 7%. This percentage is calculated by averaging over time the relative error of the non-linear error and its approximation, the linearized error.

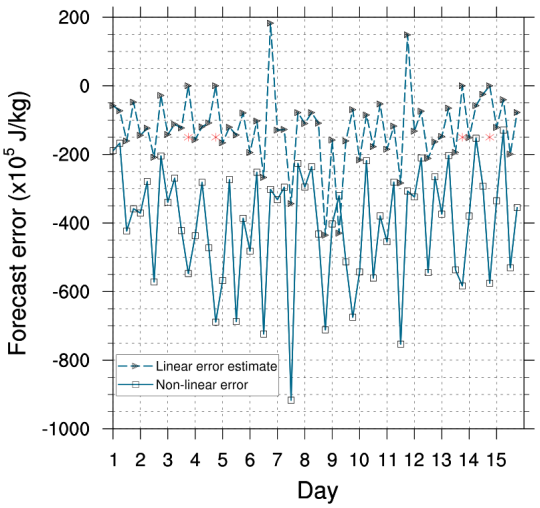


Figure 5.1: Time series of the non-linear and the approximated linear error for 01.-15.08.2014. Every day consists of four assimilation runs, thus four calculations of errors. Red asterisks denote failure of the adjoint approximation.

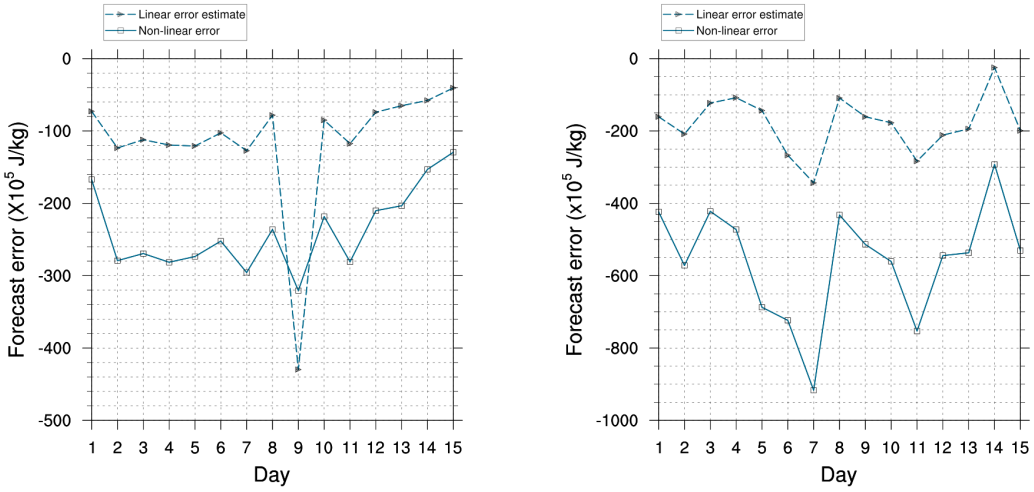
This difference might be due to the different model configuration used (e.g. different moist physics and their linearized versions), the global versus limited area configuration, the amount of data used to define the reference and the different length of forecasts examined. A relative error well below 60 % was also found in Jung et al. [2013] with the WRF model. The forecast length in that case was 6 hours and the grid spacing was 45 km. In comparison to this study, a different parameterization was used for cumulus and land-surface parameterization and this alters also the adjoint results. The chosen parameterization for this thesis follows the manual of the model (Skamarock et al. [2008]) and a finer horizontal resolution of 15km is chosen, where small scale non-linear effects may well have a larger influence. Additionally, the selected period was subject to large baroclinic developments, which might have contributed to the general poor forecast of weather centers. Nonetheless, experiments with varying forecast error approximations (see Table 2.1 in Chapter 2) do not improve the results.

It is clear that the assimilation of this observation configuration improved the forecasts. This is deduced by the negative values of the time series, meaning the forecast from the analysis is closer to the reference state than the forecast with no observations included (see Fig. 2.1). The linearized approximation follows the linear time series with the exception of 6 times, all occurring at 18 UTC on different days (days: 3, 4, 6, 11, 13 & 14). The four adjoint failure occurrences are denoted by red asterisks (days: 3, 4, 13 & 14). On days 6 & 11, big jumps to positive values are noticed (Fig. 5.1). Interestingly, when performing the same experiments for a forecast length of 12 hours, no adjoint failure is noticed but the positive spikes persisted also there.

A closer look at the experiment for the assimilation runs at 18 UTC, where the algorithm leads to inconclusive results, can be found in Appendix A. The dominant

influence of the abundant IASI data is evident in the FSO results. Especially if the reference analysis contains too little IASI data, the calculated forecast errors Eq. (2.15) and Eq. (2.16) have too small differences, which causes the failure of the algorithm. The solution would be to increase the amount of data used for the analysis. In our case, a sensor in a geostationary orbit, providing consistently data would dissolve this issue.

Obviously, the observation suite for FSO has to be carefully selected to avoid such draw-backs. In order to provide the best possible reference, as much observations as possible must be taken into account. Although some failed approximations of the linear error occur in the case study, the linearized forecast errors at 12 UTC and 18 UTC (Fig. 5.2) follow quite well the non-linear errors but always underestimating them, with the exception of 09.08 at 12 UTC.



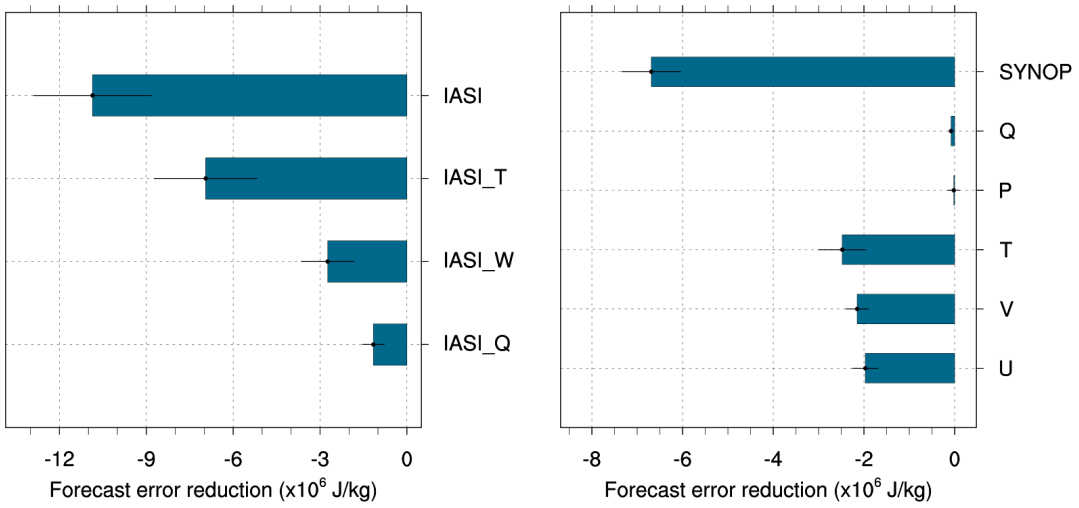
(a) Forecast error at 12 UTC

(b) Forecast error at 18 UTC

Figure 5.2: Time series of linear versus non-linear forecast error at (a) 12 UTC and (b) 18 UTC of all days between 01.-15.08.2014

5.2 IASI and SYNOP observation impact

The SYNOP and IASI observation impact on the 6-hour forecasts are calculated according to the methodology introduced in Chapter 2, Eq. (2.19). The time averaged contributions (called FECs or observation impacts) to the 6-hour forecast error reduction by assimilation of IASI and SYNOP observations are plotted in Fig. 5.3a and Fig. 5.3b, respectively. Taking into consideration the spatial distribution of observations, the time average is calculated by the summation over the domain for each assimilation time, by observation type, and then divided by the frequency of the assimilation runs. In this case, four assimilation runs per day that add up to a maximum of sixty assimilation runs in the examined time span.



(a) (b)

Figure 5.3: Time averaged error reduction by (a) IASI and (b) SYNOP observations: $IASI\_T$, $IASI\_W$ and $IASI\_Q$ stand for impact of the channels in the Temperature, Window and Water Vapor regions of the infrared spectrum, respectively. Q, P, T, V, U describe the impact of relative humidity (converted to specific humidity for the assimilation), pressure, temperature, and the two wind components. The error bars show the mean error, defined as $\sigma_{\overline{x}} = \frac{\sigma}{N}$, where σ is the standard deviation of the forecast error contributions by the IASI and SYNOP observations for the fortnight and N is the assimilation frequency.

In agreement with the previously presented error approximation (see Fig. 5.1), the negative averaged errors presented in Fig. 5.3 illustrate a reduction of the 6-hour forecast error due to assimilation. The total error reduction induced by IASI is almost twice of the total error reduction caused by SYNOP. This is related to the bulk amount of observations for this case study from IASI (around 600.000) in comparison to SYNOP (around 350.000). The numbers of all assimilated observations can be seen in Fig. 5.4. The relation of observation number and impact is obvious: especially for IASI, where more observations in the temperature region of the infrared spectrum (i.e $IASI\_T$) result in larger error reduction. This finding, is not trivial for the data assimilation since more information does not guaranty optimal initial values due to the complexity of the data assimilation systems. On the other hand, the number of observations of SYNOP measurements is approximately equally distributed between the five measured variables (Q, P, T, U, V). However, wind measurements seem to be the most valuable for the assimilation system. This is deduced by adding up the FEC by the two wind components in Fig. 5.3(b), resulting in the largest absolute number compared to the FECs by the rest of SYNOP measured variables.

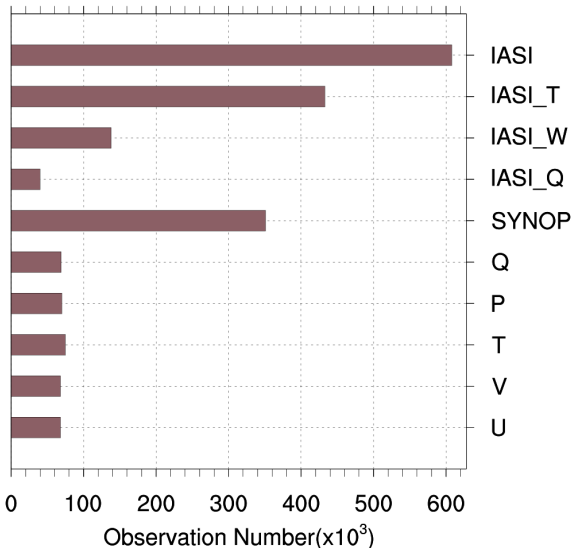
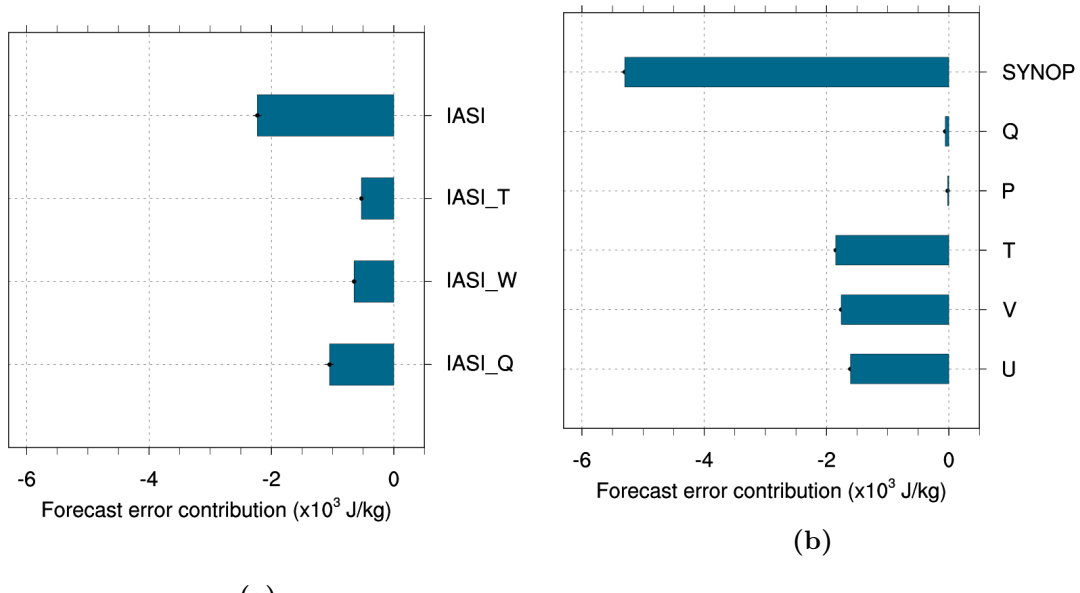


Figure 5.4: Number of observations assimilated between 01.-15.08.2014, separately depicted as IASI and SYNOP total number of observations and number of observations per measured variable and observation type.

Taking into account the number of observations for each instrument and each variable, the average forecast reduction normalized by observation number, is plotted in Fig. 5.5. Compared to the not normalized results, the qualitative contribution of SYNOP measurements is superior to IASI. It is of interest that the extra nine water vapor channels (IASI_Q) added in the assimilation for this study seem to impose more error reduction compared to the 109 channels in the temperature region of the infrared spectrum (IASI_T).

Figure 5.5: Average error reduction, normalized by observation number, of (a) IASI, where $IASI\_T$, $IASI\_W$ and $IASI\_Q$ denote the impact of the channels in Temperature, Window and Water Vapor regions of the infrared spectrum, respectively and (b) SYNOP, where Q, P, T, V, U describe the impact of relative humidity, pressure, temperature, and the two wind components.



In Fig. 5.7, the contribution of each instrument relative to the combined forecast error reduction is plotted by percentage.

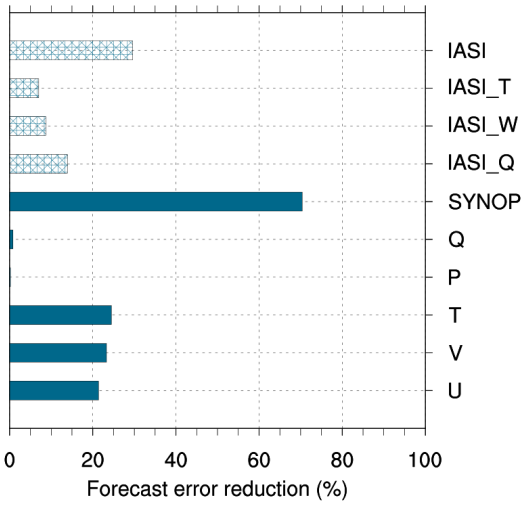
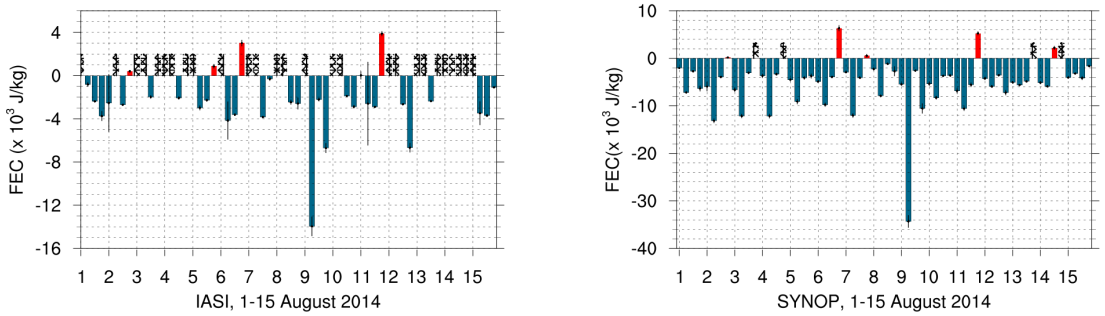


Figure 5.7: Relative forecast error reduction per instrument and variable in (%).

The percentage is calculated from the average error reduction, normalized by observation number and shown in Fig. 5.5. Almost 70% of the relative error reduction are attributed to synoptic measurements, while the remaining 30% are achieved by assimilation of IASI measurements. As underlined previously, the assimilation of nine water vapor channels had twice as much impact ($\approx$ 14%) on the forecast

error reduction compared to the temperature channels ($\approx 7\%$). The forecasts benefit more from synoptic wind measurements and not that much from synoptic relative humidity (converted to specific humidity for the assimilation) measurements.

Although the assimilation of satellite and ground based observations reduced the forecast error on average, the time series of forecast error contributions indicate days, when the forecast error increases (Fig. 5.8). More specific, positive values appear in the time series where larger positive values imply larger forecast errors by assimilation of observations. Likewise, larger absolute values on the negative range imply larger forecast error reduction by assimilation of observations.



(a) **(b)**

Figure 5.8: Time series of contributions to forecast error normalized by observation number, of (a) IASI and (b) SYNOP. FEC stands for "Forecast Error Contribution" (see Eq. (2.19)). Each day 4 assimilation runs take place, thus 4 columns per day are plotted (00, 06, 12, 18 UTC). The black crossed columns indicate that data where missing or failed the quality control. The error bars show the mean error defined as $\sigma_{\overline{x}} = \frac{\sigma}{N}$, where σ is the standard deviation and N are the number of observations.

Despite the positive columns, varying negative values per assimilation cycle dominate the time series. SYNOP data, which are more straightforward to assimilate, seem to be qualitatively more important for forecast improvement than IASI. Interestingly, both induce the largest forecast error reduction on 09.08., a day that is subject to the case of low predictability of wind and solar power forecasts, as will be seen later. The error bars show larger variability for IASI data and they are caused by the large error bars of the water vapor channels (not shown here).

Positive values appear in both time series most often on assimilation cycles at 18 UTC. In total, ten assimilation cycles are found deteriorating the forecasts and three dates out of ten coincide for IASI and SYNOP data (namely 02., 06. and 11.08, at 18 UTC). The largest degradation of the forecasts due to assimilation is found on 06.08 at 18 UTC for SYNOP data and on 11.08 at 18 UTC for IASI data. These are also the dates with unexpected positive spikes in the time series of linear versus non-linear forecast error. With the exception of these two dates, the largest degradation from SYNOP assimilation is found on 14.08 at 12 UTC. On this day, the assimilation of the pressure measurements caused the largest forecast deterioration in comparison to the other measurements by SYNOP stations. The positions of the stations with the largest positive values can be tracked down as plotted in Fig. 5.9.

It shows the forecast error contribution for each SYNOP station providing pressure observations for the assimilation on 14.08 12 UTC. A large number of stations that cause forecast degradation is placed in the region $45°N$ to $60°N$ and $20°E$ to $30°E$. In total, more than 50% of the stations in the domain degrade the forecast due to assimilation of their measured pressure components. This worst performance of SYNOP pressure measurements in the fortnight simply states that the forecasted pressures, initiated from the analysis, had the largest difference from the reference run pressures, compared to the rest of assimilation cycles in the fortnight. Subsequently the corrections applied to the analysis on 14.08 at 06 UTC by SYNOP were not lost but they rather declined the quality of the forecast. It is unclear whether this degradation is due to bad observations, false forecast or even untrustworthy reference state. Unfortunately FSO can not identify the reason but rather alarm that this set-up, with this assimilation and forecast model and these observations is degrading the 6-hour weather forecast.

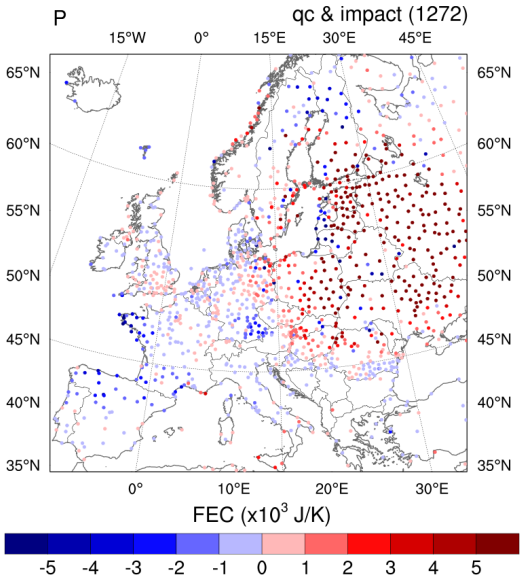
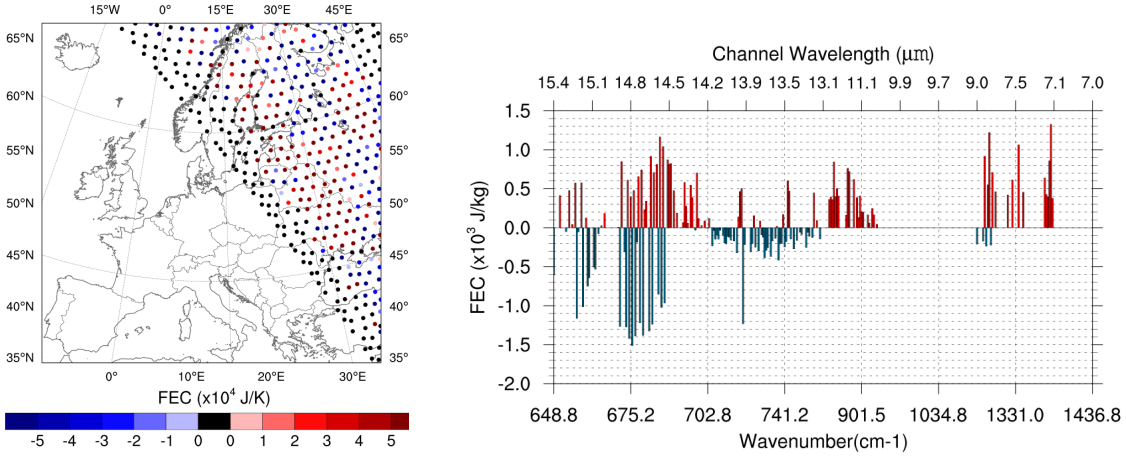


Figure 5.9: Impact of SYNOP pressure observations assimilated on 14.08, 12 UTC on the 6 hour forecast."FEC"stands for Forecast Error Contribution and "qc & impact" refers to the number of stations across the simulation domain that successfully passed the quality control and had an impact on the forecast.

The worst performance of IASI on 05.08 at 18 UTC (excluding 11.08, at 18 UTC) is illustratively demonstrated in Fig. 5.10, which shows the impact of IASI observations on the forecast errors at this assimilation time. Positive values, thus forecast deterioration, dominate the simulation domain over Europe (Fig. 5.10a). The pixels beneficial to the forecast are not sufficient to compensate the overall bad results. This becomes evident in Fig. 5.10b, where the impact of the individual channels is plotted, normalized per pixel number. Out of the 158 channels, 53% degrade the forecast, whereof 28% are caused by assimilation of temperature channels.



(a) **(b)**

Figure 5.10: (a) Impact of assimilation of IASI pixels on the 6-hour forecast on 05.08, at 18 UTC. For each pixel maximal 158 channels are assimilated. Black pixels correspond to quality control failure, which in this case is mostly due to observations retrieved from large scan angles. Positive values, indicated in red, describe forecast degradation (Fig. 5.10). Pixels improving the 6-hour forecast are in the blue range of colors. (b) Impact of IASI channels normalized by pixel number.

The monitoring aspect of FSO is emphasized in the presentation of the aggregated results. Observations that potentially degrade the forecast can be excluded from the assimilation procedure and locations can be classified to be characterized by high or low impact. Furthermore, potential applications of FSO become apparent like tracking systematic errors related either to weather conditions or to the model's ability to properly assimilate or even persistent technical malfunctions of instruments. This can be realized if FSO is set-up on an operational basis, as a computationally cheap alarm. Consequently, with a specific model and observation set-up, a value of FEC can be defined indicating even extreme forecast error events. In those cases, the observation data set and assimilation system can be examined, in terms of observation error tuning for the observations case or the model set-up for the assimilation case. Of course it is of interest to track the weather circulation on those situations, in case a specific pattern arises that is ill-represented by the model or a more dense observation network should be considered.

5.2.1 Observation System Experiments

For the verification of the results, OSEs are utilized. These experiments, include assimilation and forecast cycles under observation denial scenarios (Dumelow [2003]). The reference run includes the 'all data' scenario and statistical scores such as the root mean square error are usually calculated. To make the OSEs results comparable to FSO results, the forecast error is defined in terms of energy norm (see Eq. (2.1)). The reference scenario is defined as,

$$\mathbf{e}_{\mathbf{ref}}^{\mathbf{t_0+\Delta t}} = (\mathbf{x}_{\mathbf{a}}^{(\mathbf{t_0},\mathbf{t_0+\Delta t})} - \mathbf{x}_{\mathbf{t}}^{\mathbf{t_0+\Delta t}})^{\mathbf{T}}\mathbf{C}(\mathbf{x}_{\mathbf{a}}^{(\mathbf{t_0},\mathbf{t_0+\Delta t})} - \mathbf{x}_{\mathbf{t}}^{\mathbf{t_0+\Delta t}}) \tag{5.1}$$

where $x_a^{(t_0,t_{0+\Delta t})}$, is a forecast initiated from an analysis at time t_0 that includes all observations. To be consistent with the FSO configuration, the "all data" scenario includes the same observation suite as defined in table (Table 4.4). As in FSO, x_t is an analysis at verification time $\mathbf{t_0} + \mathbf{\Delta t}$ that includes all observations. The forecast errors from the IASI and SYNOP observation denial experiments,

$$\mathbf{e}_{\mathbf{iasi}}^{\mathbf{t_0+\Delta t}} = (\mathbf{x}_{\mathbf{a},\text{no iasi}}^{(\mathbf{t_0},\mathbf{t_0+\Delta t})} - \mathbf{x}_{\mathbf{t}}^{\mathbf{t_0+\Delta t}})^{\mathbf{T}}\mathbf{C}(\mathbf{x}_{\mathbf{a},\text{no iasi}}^{(\mathbf{t_0},\mathbf{t_0+\Delta t})} - \mathbf{x}_{\mathbf{t}}^{\mathbf{t_0+\Delta t}}) \tag{5.2}$$

$$\mathbf{e}_{\mathbf{synop}}^{\mathbf{t_0+\Delta t}} = (\mathbf{x}_{\mathbf{a},\text{no synop}}^{(\mathbf{t_0},\mathbf{t_0+\Delta t})} - \mathbf{x}_{\mathbf{t}}^{\mathbf{t_0+\Delta t}})^{\mathbf{T}}\mathbf{C}(\mathbf{x}_{\mathbf{a},\text{no synop}}^{(\mathbf{t_0},\mathbf{t_0+\Delta t})} - \mathbf{x}_{\mathbf{t}}^{\mathbf{t_0+\Delta t}}) \tag{5.3}$$

are then compared to the reference forecast error.

In Fig. 5.11 the temporally averaged 6-hour forecast errors for each assimilation time (at 00, 06, 12, 18 UTC) are plotted. When both data sets are included in the assimilation cycles, the forecast error is on average reduced as is seen in the 'All Data' columns. The total forecast error appears to increase with the forecast hour up to 18 UTC, with the absence of IASI data increasing the forecast error more than the absence of SYNOP data. The bulk number of IASI data available for assimilation at 18 UTC have a drastic impact on error reduction. This is deduced by the large forecast error measured when they are absent.

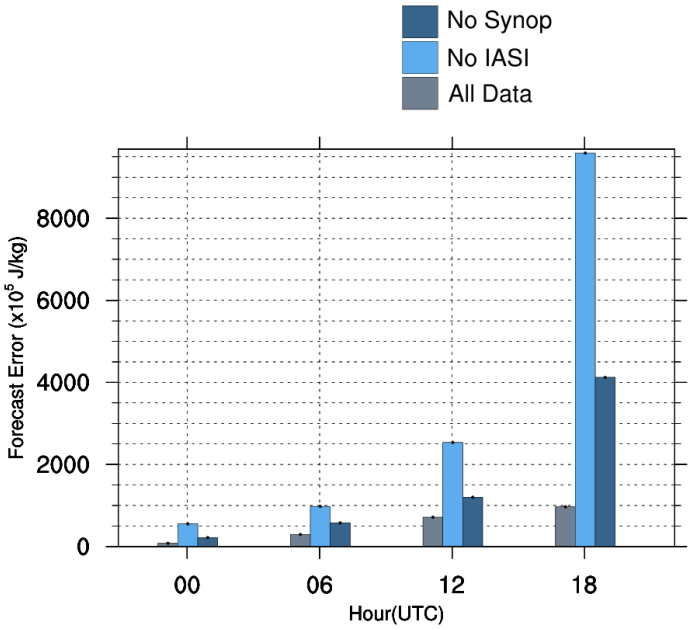


Figure 5.11: The averaged forecast error in terms of the energy norm is calculated for the four assimilation hours over the time span 01.-15.08.2014. The forecast error is calculated for the reference (Ref-gray column bar) that includes all observations, the No IASI (light blue column bar) and the No Synop (dark blue column bar) experiments.

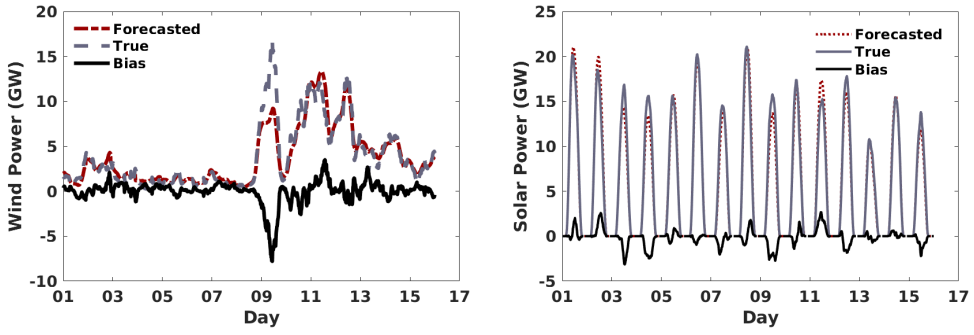
Qualitatively the OSE results support the findings with FSO. This applies to the non-normalized results, where the bulk IASI data have the largest contribution to the forecast error reduction in the fortnight. In support of this, OSE results show that the absence of IASI data increases more the forecast error than the absence of SYNOP data. One exception to the relation of observation impact with the number of observations does exist at 06 UTC. According to Fig. 4.9, the percentage of assimilated SYNOP observations at 06 UTC is higher than the percentage of IASI assimilated observations at 06 UTC. The OSE results, in contrast, show a larger increase of forecast errors when the IASI data are absent at 06 UTC compared to the forecast error when SYNOP data are absent (Fig. 5.11). Despite this, the time series of the FEC normalized by observation number showed larger forecast error reduction by SYNOP data on all assimilation runs, including 06 UTC.

5.3 Poor predictability cases

To apply FSO for energy meteorology, erroneous solar and wind power events are sought for which the weather pattern is examined and the contribution of the observation network is evaluated through the applied algorithm. In this chapter those days are identified based on forecasted and actual total power values produced from German power plants. The accurate prediction of large fluctuations (ramp events) of solar and wind power is very important for the integration of renewable energies in the electrical power grid (Bueno et al. [2017]). Ramp events can take place at different space and time scales and in both positive and negative directions (Florita et al. [2013]). Wind power ramp events are caused by atmospheric phenomena like wind gusts, thunderstorms, cyclones and low-level-jets (Freedman et al. [2008]). Solar power ramp events, occur due to diurnal variability by clouds, mineral dust outbreaks, biomass burning plumes or PV panes blurred by deposited snow or dust layers (Cui et al. [2017]).

In order to identify these prediction error events, data available by Germany's Transmission System Operators (TSOs) are examined. The data set of measured and forecasted power values (obtained by TSO-50hertz [2014]) has a temporal resolution of 15 minutes and the day-ahead power forecasts are weighted averages of various NWP models. The examined time frame is the first fortnight of August 2014 and includes the extreme wind power forecast error event on 09.08., as identified by Good [2017]. A power output model using meteorological input from WRF would have been a better choice to identify the erroneous days and use FSO to improve the observation network but unfortunately it was not available for this study.

The time series of the first fortnight of August 2014 of the forecasted versus the real power feed-in for the wind and solar power (in GW) for Germany is plotted in Fig. 5.12. For the wind power time series (Fig. 5.12a), the second half of the fortnight is apparently of more interest with the notable under-prediction of wind power on 09.08.. In contrast to this error event, almost 65% of the wind power forecasts were overestimating the actual power output. In the solar power time series, a persisting under-prediction of the daily maximum actual solar power output can be noticed (Fig. 5.12b).



(a) (b)

Figure 5.12: Time series (01.-15.08.2014) of day-ahead forecasts and actual total power values for (a) wind and (b) solar from German power plants. The bias is depicted by the black solid line.

Keeping in mind that the configuration of the FSO experiment consists of assimilation and 6-hour forecast cycles at 00, 06, 12 and 18 UTC, it is of interest to examine the time series at these hours during the fifteen days experiment (Fig. 5.13). It was found that the largest differences between forecasted and actual power values for both wind and solar occurred at 06 and 12 UTC.

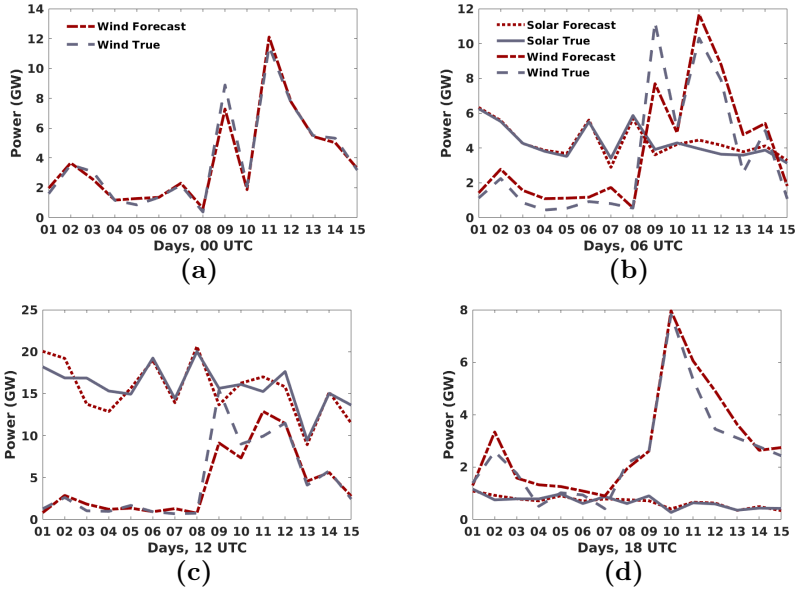


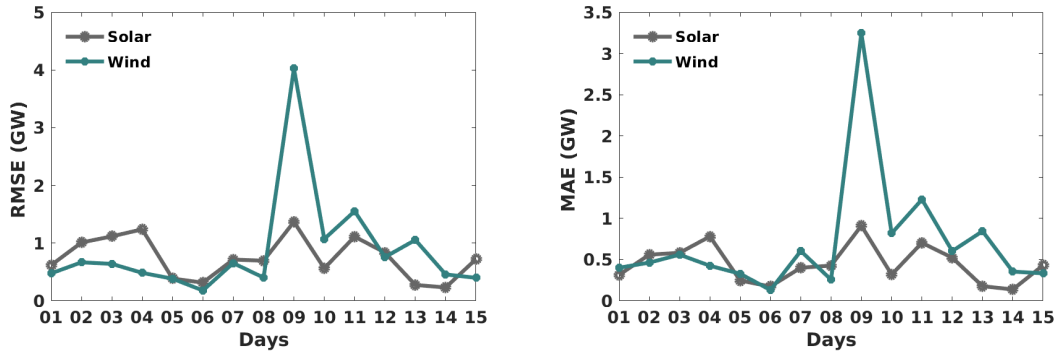
Figure 5.13: Time series (01.-15.08.2014) of day-ahead forecasts and actual total power values for wind and solar power for Germany at (a) 00 UTC, (b) 06 UTC, (c) 12 and (d) 18 UTC. All plots share the same legend.

To identify the erroneous forecast days, statistical measures such as the RMSE and the Mean Absolute Error (MAE) are employed separately for solar and wind power. The RMSE, aggregates the magnitude of the daily error of the predictions in one value for each day. It is calculated according to Eq. (5.4), where x_f and x_t refer to forecasted and measured power values, respectively. MAE (Eq. (5.5)) provides the average absolute values of the errors. In contrast to RMSE, it is proportionally influenced by each error.

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{\mathbf{n=1}}^{\mathbf{N}}(\mathbf{x_f}-\mathbf{x_t})^{\mathbf{2}}}{\mathbf{N}}} \tag{5.4}$$

$$\mathrm{MAE}=\frac{\sum_{\mathbf{n=1}}^{\mathbf{N}}|\mathbf{x_f}-\mathbf{x_t}|}{\mathbf{N}} \tag{5.5}$$

The time series of both measures are plotted in Section 5.3. The large wind power errors stand out in the second half of the fortnight, with maximum error occurring on 09.08.2014. No extreme forecast error is noticed for solar power, although the largest error occurs also on 09.08.. In Table 5.1 the most erroneous dates for wind and solar power according to the statistical measures are listed. In the the study period the largest error events in terms of power discrepancies between actual and foretasted power (GW) are found for wind power.



(a) (b)

Figure 5.14: Time series (01.-15.08.2014) of daily (a) RMSE and (b) MAE for wind and solar power for Germany.

Table 5.1: Dates with highest RMSE and MAE scores for wind and solar power.

Solar			
Dates (MM.DD)	RMSE (GW)	Dates (MM.DD)	MAE (GW)
09.08	1.37	09.08	0.91
04.08	1.23	04.08	0.78
03.08	1.11	11.08	0.70
Wind			
09.08	4.03	09.08	3.24
11.08	1.55	11.08	1.22
10.08	1.10	13.08	0.84

The hourly RMSE for solar and wind power in the examined fifteen days is plotted in Fig. 5.15. As was already noticed in Fig. 5.13 and confirmed in the next plot (Fig. 5.15), 06 and 12 UTC have the largest RMSE for both wind and solar power. Higher wind error scores persist over several hours in contrast to solar power prediction errors.

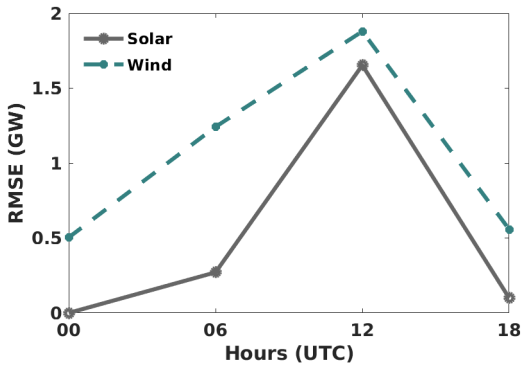


Figure 5.15: RMSE per hour (00, 06, 12, 18 UTC) for 01.-15.082014.

For the listed days in Table 5.1, the time series of forecasted and actual total power values are plotted in Fig. 5.16 and Fig. 5.17. During the days with the largest

wind error events, the TSOs' day-ahead forecasts were able to identify the majority of the ramps but failed to estimate the magnitude of the power fluctuations. On the other side, the TSOs' solar forecasts on the most erroneous days did not only miss many ramp events but also misplaced them in time (Fig. 5.17a). This can be understood, as the effect of transient clouds is typically difficult to predict by the weather models.

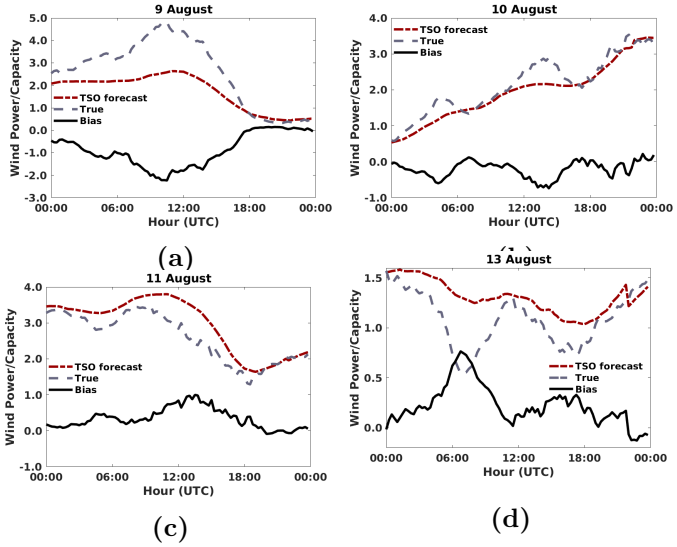


Figure 5.16: Time series of days with large error events (see Table 5.1) for wind power over Germany. The total power values are normalized by the installed capacity of 35 GW for Germany on 2014.

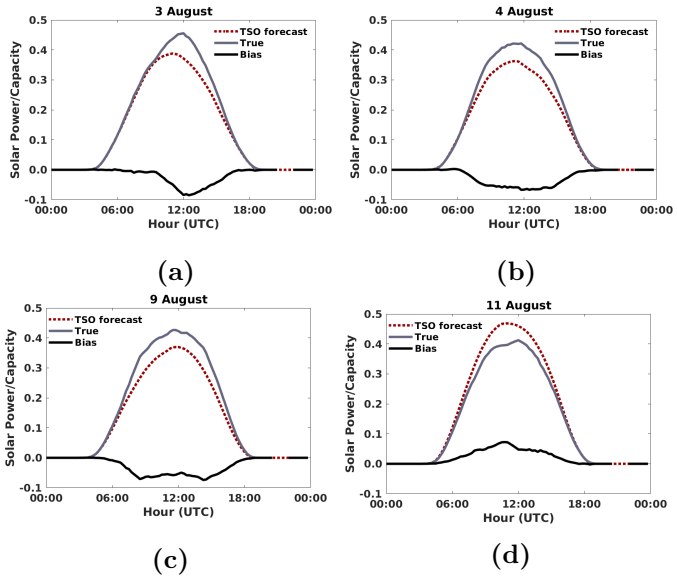


Figure 5.17: Time series of days with large error events (see Table 5.1) for solar power over Germany. The total power values are normalized by the installed capacity of 39 GW for Germany on 2014.

To comprehend the forecast errors quantitatively the total forecasted and measured energy of these days was calculated. The integrals over time of forecasted and measured power were approximated numerically by applying the trapezoidal rule (with a 15 minutes time interval). The relative error (in %) of the forecasts, on the days with the largest errors, is presented in Table 5.2. It is calculated in percent by $((x_t - x_f)/x_t) * 100$ where x_t and x_f refer to the daily measured and forecasted energy, respectively. Positive relative error values imply under-prediction of the energy produced that day, by the TSOs' day-ahead forecasts and negative values imply over-prediction. The solar energy forecasts underestimated the actual energy by a range of 10-16 % with an exception of over-prediction on 11.08.. Wind energy forecasts, have a larger range of over/under prediction on the erroneous days, with an astonishing relative error of 34% on the exceptional wind error event.

Table 5.2: Relative forecast error in percentage with respect to real energy production for days with the largest RMSE and MAE scores.

Solar		**Wind**	
Dates (MM/DD)	Relative error (%)	Dates (MM/DD)	Relative error (%)
09.08	16	09.08	34
04.08	14	11.08	-13
03.08	10	10.08	10
11.08	-13	13.08	-21

In the following Section 5.3.1, the weather situation on the poor solar and wind power predictability days is examined in order to identify potential patterns that might occur. Then, the described observation configuration is evaluated for these days with the help of FSO in Section 5.4.

5.3.1 Weather conditions

In this section the weather situation of the low solar and wind power predictability days is analyzed. The geopotential height distribution at the 500hPa surface is plotted together with the sea level pressure contours in Fig. 5.18. These snapshots of the atmosphere were plotted from the 3DVAR analyses of the FSO experiment, where the background information originates from model forecasts initiated 6-hours before the analyses. The dates of plotted analyses are chosen according to the largest discrepancies between predicted and actual solar and wind power (see Section 5.3, Fig. 5.16 and Fig. 5.17).

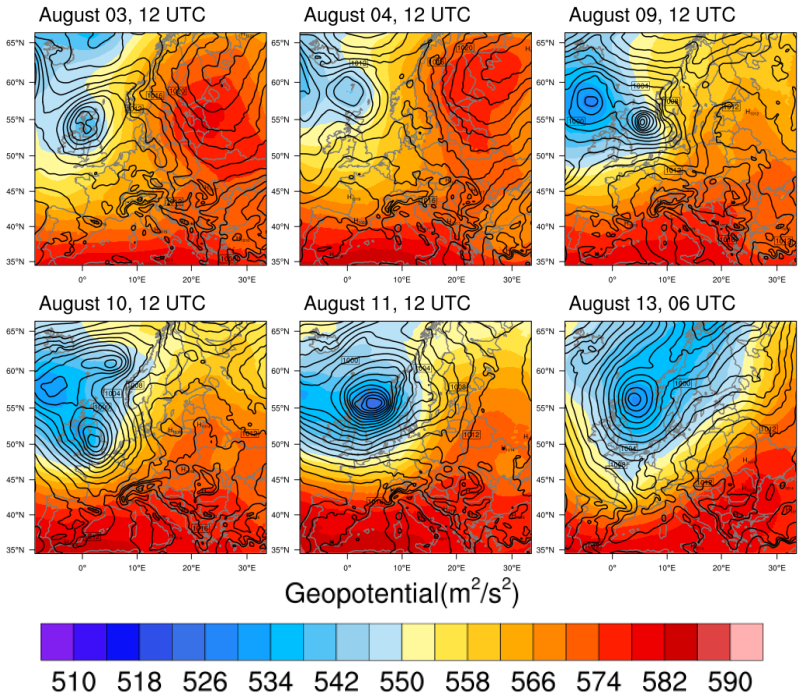


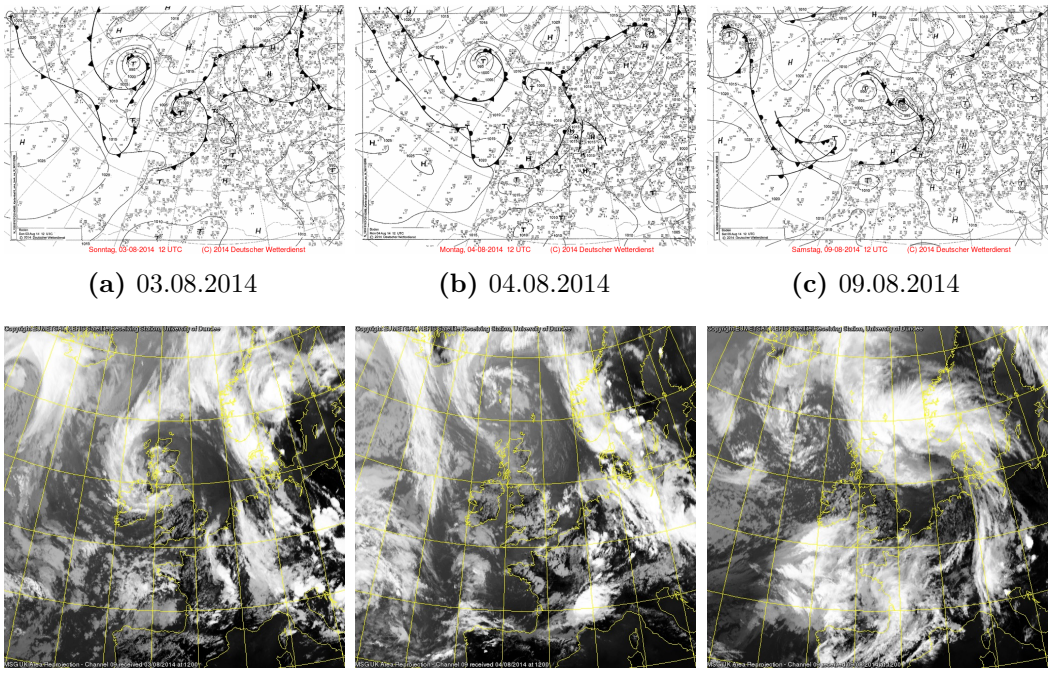
Figure 5.18: Geopotential height in decametres at 500 hPa (color-scale) and sea level pressure (black contours) for days with large wind (09., 10., 11., 13.08.2014) and solar power (03., 04., 09. and 11.08.2014) errors.

Isobaric (constant pressure) maps of the upper atmosphere are valuable to illustrate the movement of surface pressure systems. For example, developing surface storm systems are deep lows that intensify with height. Thereupon, a surface low-pressure area can appear on an upper-level chart as either a closed low or a trough (Ahrens [2011]). They steer the surface low-pressure systems and bring cold air masses from the pole towards the European continent. This is especially prominent in the low wind power predictability cases on 11.08., where a sharp pressure gradient indicated strong southerly winds along the British Isles and on 09.08., at the cut-off low, with prevailing southerly winds along the North Sea.

To fully comprehend the weather associated with these lows, that vary in intensity and extent, the surface analyses and the imagery of the cloud scenes over Europe is given in Fig. 5.19 and Fig. 5.20. The surface analyses are provided by the DWD. The imagery from SEVIRI at the thermal infrared band (10.8 μm) records the thermal emission from surface and from the top of the clouds (Petty [2004]).

The frontal wave that gave birth to the cyclonic circulation on 03.08. over the UK, formed on 02.08. around 06 UTC. On 03.08. at 12 UTC the system had already build up to a fully developed open wave over Scotland (Fig. 5.19a), moving eastward. On 04.08., 12 UTC the dissipation of the cyclone, positioned now north of Scotland, had begun (Fig. 5.19b). Taking a look at Fig. 5.19d and Fig. 5.19e sharper differences at the center of the low between the brightness temperatures of the clouds is noticed in the first figure, whereas in the second more shallow clouds appear. Scattered clouds appear in western Germany in both images and larger patches of high clouds in eastern Germany. On those days, Germany's TSOs under-predicted the total solar energy production by $\sim 15\%$ (Table 5.2).

Coldest brightness temperatures are highlighted with white and warmer brightness temperatures with gray or black. The contrast between the deep thunderstorms (depicted with white) and the surrounding lower and thus warmer cloud tops are prominent over all examined days (Lower panels in Fig. 5.19 and Fig. 5.20). In the following paragraphs, isobaric upper level maps, surface analyses and cloud imagery will be used to describe the weather circulation on low predictability days of, first solar, then wind, and last, both solar and wind power output.

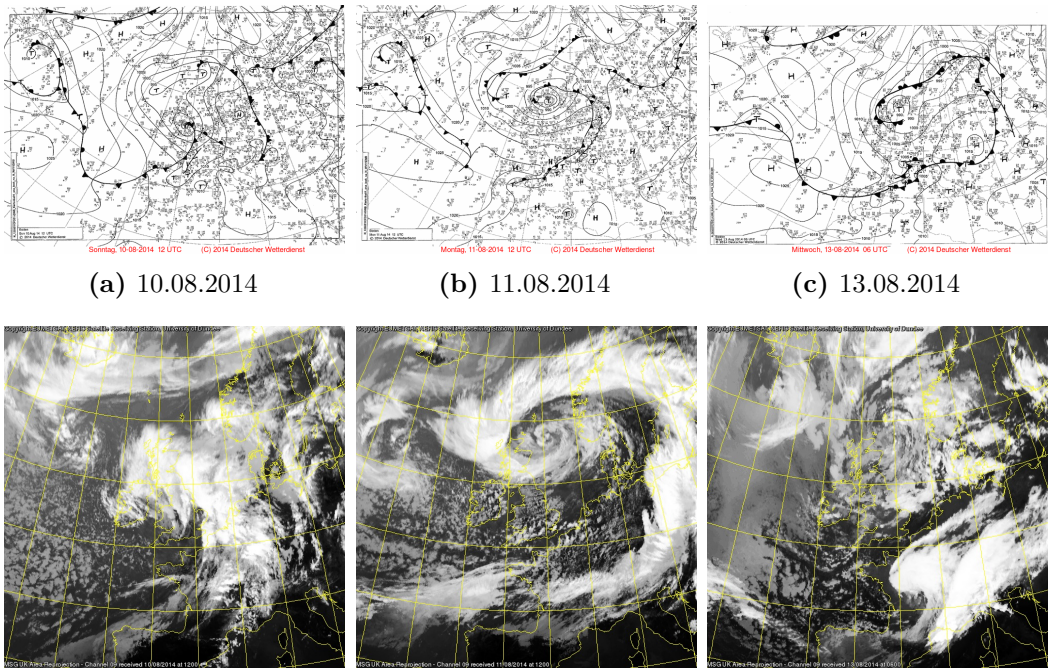


(a) 03.08.2014 **(b)** 04.08.2014 **(c)** 09.08.2014

(d) 03.08.2014 **(e)** 04.08.2014 **(f)** 09.08.2014

Figure 5.19: Upper panel: surface analyses by the German weather service for (a) 03.08.2014, 12 UTC (b) 04.08.2014, 12 UTC (c) 09.08.2014, 12 UTC. Black, thin contour lines refer to the sea level pressure and black thick lines with triangles and half-circles describe cold and warm frontal position, respectively. **Lower panel:** Meteosat Second Generation satellite, imagery by the SEVIRI (UK Projection), channel 9 at 10.8 μm in the thermal infrared band.
Source (a-c): $http://www1.wetter3.de/archiv\_gfs\_en.html$
Source (d-f): $http://www.sat.dundee.ac.uk$

The weather patterns on the low wind power predictability days on 10 and 13.08. are depicted in Fig. 5.20a and Fig. 5.20c. Starting with 10.08., a cold front was moving south-eastwards across France and Germany (Fig. 5.20a) and was steered by the deep low-pressure system situated over the UK. This low was part of a family of lows across the trough in Fig. 5.18 (10.08. at 12 UTC). The steep pressure gradient that produced strong south-west surface winds along the UK is also clearly visible. Germany was almost completely covered with clouds, with the exception of a clear sky stripe in eastern Germany (Fig. 5.20d). On 13.08., an expanded cold air mass in the upper atmosphere over Europe (Fig. 5.18 (13.08. at 06 UTC)) produced a broad surface low-pressure system situated over the North Sea and secondary lows over the central Europe (Fig. 5.20c). At the latter ones, high clouds indicating deep thunderstorms can be identified by the bright white color in Fig. 5.20f. A 10 % under-prediction of the total daily wind energy production by the forecasts was found for 10.08. and an over-prediction by 21% on 13.08. (Table 5.2).



(a) 10.08.2014 **(b)** 11.08.2014 **(c)** 13.08.2014

(d) 10.08.2014 **(e)** 11.08.2014 **(f)** 13.08.2014

Figure 5.20: Upper panel: surface analyses by the German weather service for (a) 10.08.2014, 12 UTC (b) 11.08.2014, 12 UTC (c) 13.08.2014, 06 UTC. Black, thin contour lines refer to the sea level pressure and black thick lines with triangles and half-circles describe cold and warm frontal position, respectively. **Lower panel:** Meteosat Second Generation satellite, imagery by the SEVIRI (UK Projection), channel 9 at 10.8 μm in the thermal infrared band.
Source (a-c): $http://www1.wetter3.de/archiv\_gfs\_en.html$
Source (d-f): $http://www.sat.dundee.ac.uk$

The last days examined are 09. and 11.08.. The highest daily RMSE of forecasted total power values of Germany, over the examined period, was found for both solar and wind power on 09.08. On the other hand 11.08 was the second most erroneous wind power forecast day, whereas it was found on the fourth place (descending RMSE order) for the solar power forecast cases in the examined two weeks. On both days, a surface low-pressure system was situated between northern England and Norway. On 11.08. (Fig. 5.20(b)), a cold front extended all the way down to south Europe in a pronounced spiral cloud pattern and scattered clouds persisted in central Europe (Fig. 5.20(e)). On 09.08., the cyclone over the North sea produced the familiar comma shaped cloud pattern (Fig. 5.19(f)) as a result of the occluded front (Fig. 5.19(c)) that is accompanied by the most intense weather in a life of a cyclone. Eastern Germany was affected by this surface low-pressure system and the western part was affected by scattered clouds by the low-pressure system situated over North-western France. On both days, steep pressure gradients visible in Fig. 5.18[09.08, 12 UTC], indicate strong south-westerly winds over Germany. The over-prediction of wind energy production for Germany on 11.08. was $\sim$ 13%. In contrast, on 09.08. the TSOs under-predicted the wind power by $\sim$ 34%.

Clearly, the ability to precisely forecast the position of the surface low-pressure

systems along with the magnitude of the baroclinic instability comes in question. On 09.08., the NWPs under-predicted the magnitude of the instability as stronger than expected winds occurred producing higher wind power feed in than forecasted. This instability evolved in the development of a secondary low pressure system (see Fig. 5.18, 09.08. at 12 UTC). The significantly mispredicted surface winds, produced an extreme forecast error case (Berndt [2018]). In this case, the TSOs were anticipating less power feed in for the grid than what was actually received leading to one of the most significant losses of revenues in the year. The evaluation of the configuration of the SYNOP and IASI observation network, based on the FSO results, follows in Section 5.4.

5.4 Evaluation of the results

In the following, the results of the FSO algorithm are examined on the days of low wind and solar power prediction. First the effect of assimilation of the IASI and SYNOP observations is examined (Section 5.4.1) in terms of distance of the observations from the background model information and the analyses. Furthermore, the most valuable SYNOP parameters and IASI group channels are identified in section Section 5.4.2. The impact is then attributed to the individual SYNOP stations and IASI pixels, revealing locations that improve or occasionally degrade the intraday forecast error. The impact per IASI channel is examined and validated in terms of effective improvements. In section Section 5.4.3, the range of the observation impact is explained through the relationship to the innovation vector. Larger distance of the observations from the model background is linked to larger observation impact. The model's over/under-prediction of the observations is examined with respect to the observation impact.

5.4.1 Assimilation results

In order to perceive the effect of IASI and SYNOP observations in the assimilation scheme from different perspectives the following graphs are presented. The examined analyses are the ones for which the observation impact on the 6-hour forecast is analyzed in Section 5.4.2. In Fig. 5.21, the daily mean and standard deviation of Observation-Minus-Background (OMB) and Observation-Minus-Analysis (OMA) for the five SYNOP variables are plotted. The two wind components are plotted in Fig. 5.21a. The same plots are produced for the IASI channels on the same dates, where the average for each channel is calculated according to the number of pixels (Fig. 5.22, Fig. 5.23). The OMB, or innovation vector (Eq. (3.2)), reveals whether the model over-predicts (negative OMB) or under-predicts (positive OMB) the observations. The OMA reveals the tendency of the assimilation algorithm to tend towards the background information or the observation. This preference depends on the observation and forecast covariance matrix errors. Small OMA values, mean a preference to observations and larger values indicate a preference to the background model information.
Larger OMB values in comparison to OMA values, indicate that the analysis has larger values than the backgound due to assimilation of observations. Smaller OMB values in comparison to OMA values, imply that the analysis has smaller values than the backgound, due to the assimilation of observations. When OMA and OMB are very close, small corrections to the backgound are attributed to the observations.

Starting with the SYNOP wind variables in Fig. 5.21a, on all erroneous wind power dates (09., 10., 11. and 13.08.) the model was over-predicting the wind velocity. More specific, the departures (in absolute values) of the observations from the model's background are found to be larger for the meridional (V) than the zonal (U) wind.

The smaller OMA in absolute values, implies that the algorithm, shifted the analysis closer to the observation value. These corrections by the assimilation algorithm brought the assimilation of meridional wind component on the highest impact

ranking list, as will be shown later. The notable under-prediction of temperature by the model on 09.08. and 11.08. at 06 UTC and the subsequent correction by the assimilation system, brought the SYNOP temperatures on the highest ranking position on those two days (Section 5.4.2). As seen in Section 5.3.1, a surface low pressure system situated between northern England and Norway was the dominant feature on those two days with cold fronts sweeping across Europe (Fig. 5.19(c) and Fig. 5.20(b)). The changes of temperature that accompany cold fronts and the ability of the model to predict them, indicated that the introduction of temperature information in the model was the most beneficial information to improve the short-term forecast related to this weather circulation.

For the IASI channels, the model is in majority overestimating the IASI_W channel data and along with IASI_Q, they appear to have the largest standard deviations. The IASI_T channels are constantly under-predicted by the model and the smallest values of OMA and OMB are found in the temperature channel range $14.5 - 14.0\mu m$. These channels were found to also have the smallest impact on forecast error reduction on the examined low predictability days, as will be seen later. The nine additionally included IASI_Q channels, are overestimated by the model and brought closer to the observations, as seen by the smaller OMA compared to the OMB.

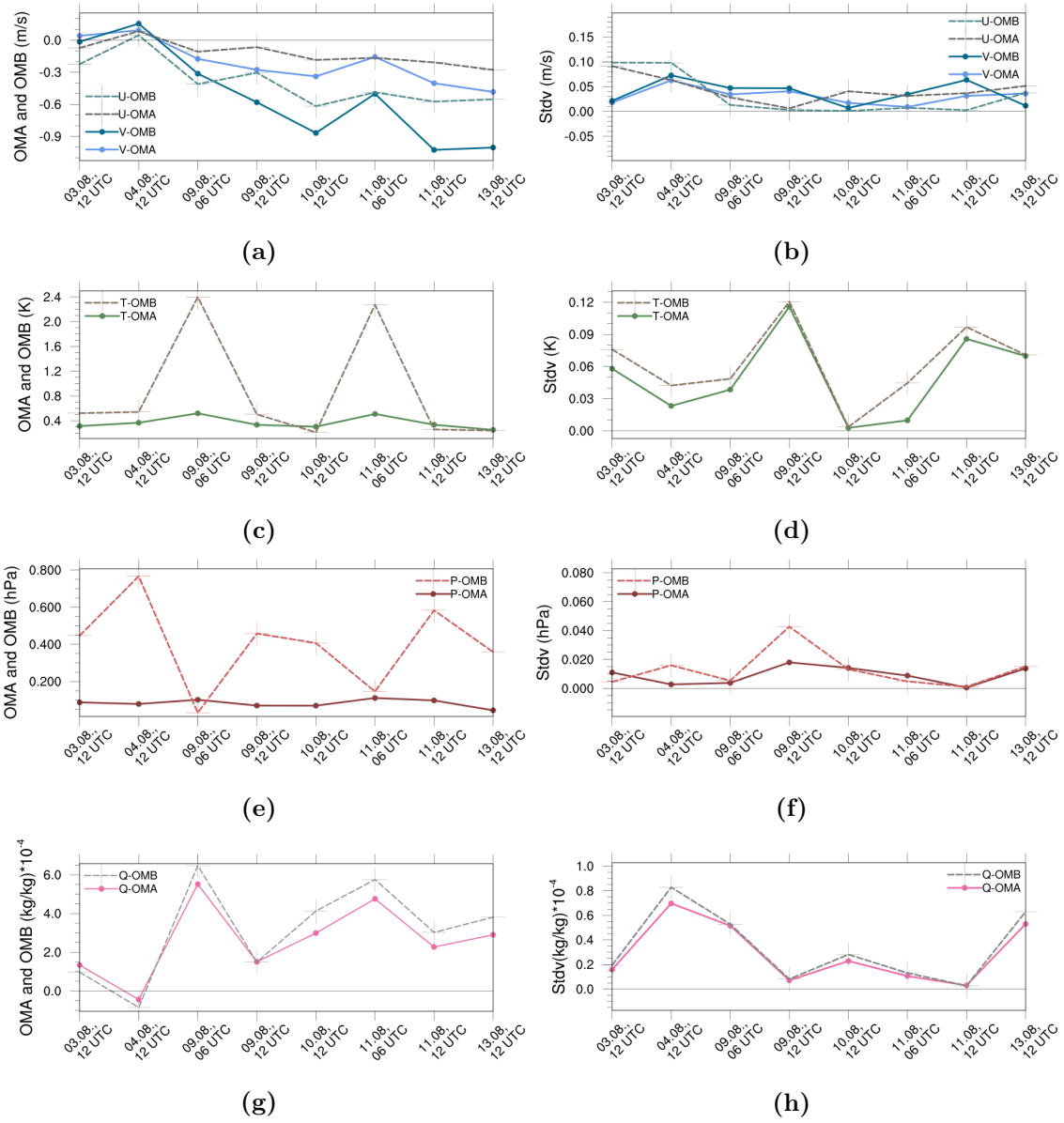
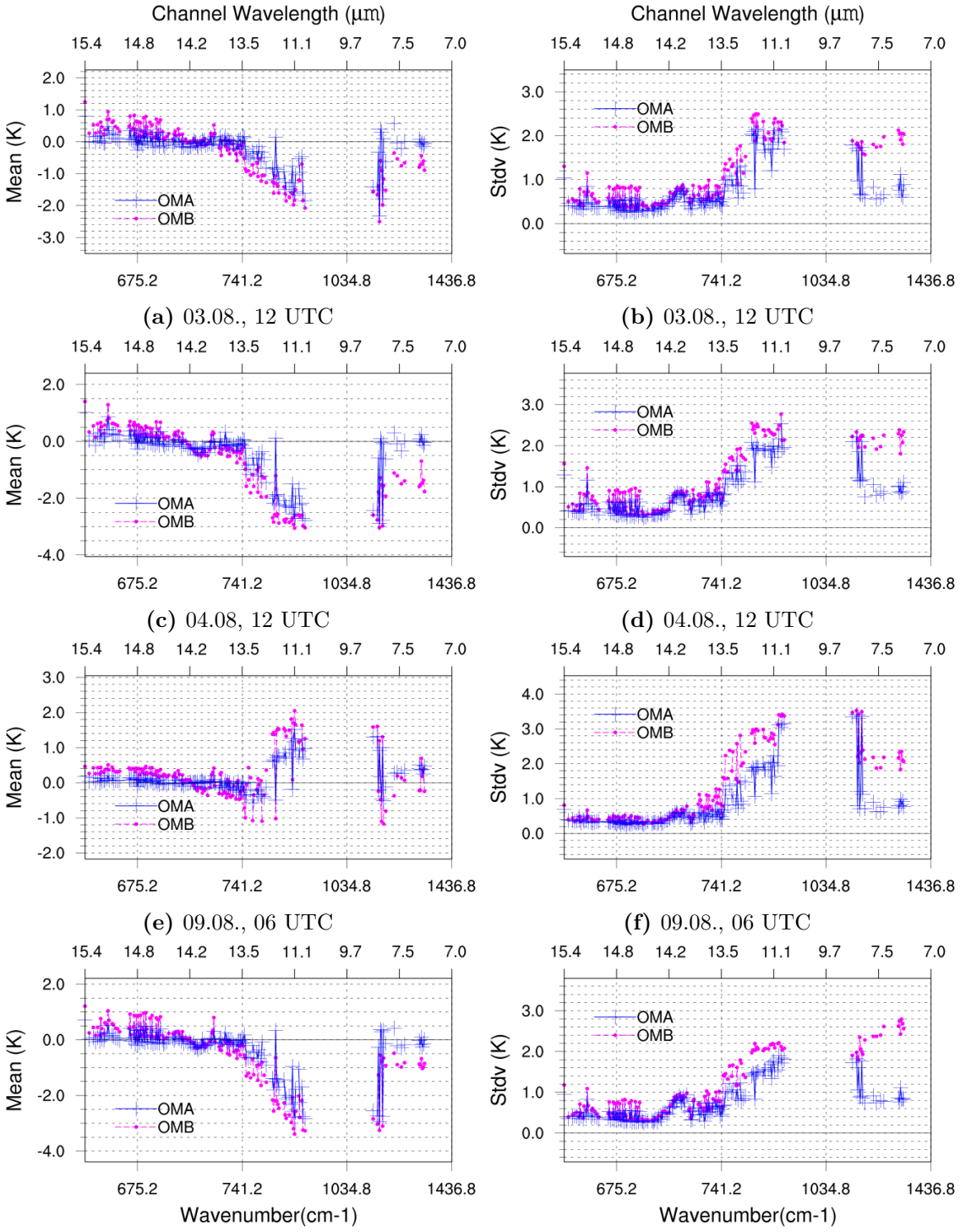


Figure 5.21: Time series of mean (left) and standard deviation (right) of the OMA and OMB for the SYNOP (a,b) U and V wind components, (c,d) temperature (e,f) pressure and (g,h) specific humidity. The dates on the x-axis are the analyses dates examined later on for the low predictability cases.



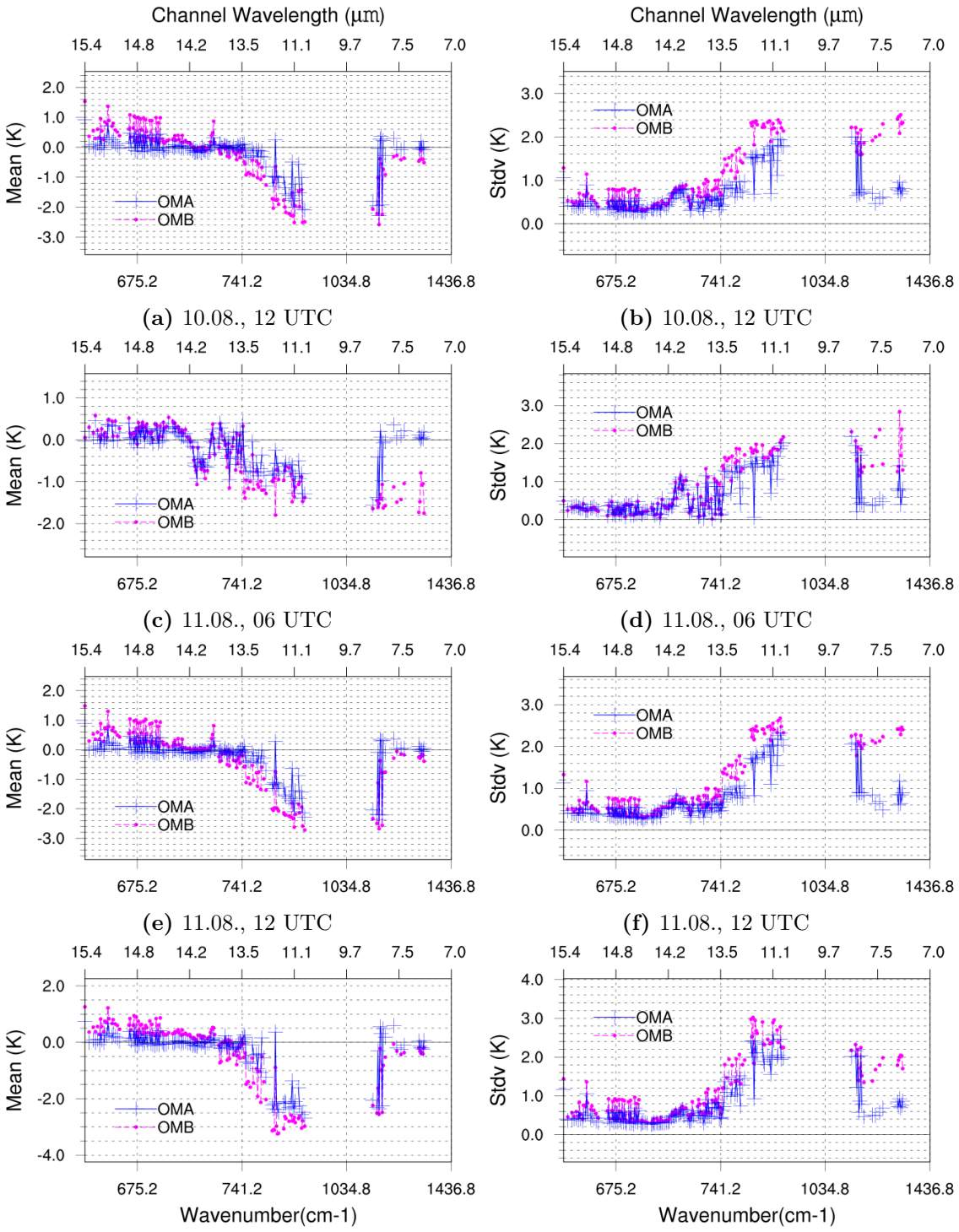
(a) 03.08., 12 UTC **(b)** 03.08., 12 UTC

(c) 04.08, 12 UTC **(d)** 04.08., 12 UTC

(e) 09.08., 06 UTC **(f)** 09.08., 06 UTC

(g) 09.08., 12 UTC **(h)** 09.08., 12 UTC

Figure 5.22: Analyses from 04.08., 12 UTC to 09.08., 12 UTC: Mean (left) and standard deviation (right) OMA and OMB for the assimilated IASI channels. The average for each channel is calculated according to the number of pixels.



(a) 10.08., 12 UTC **(b)** 10.08., 12 UTC

(c) 11.08., 06 UTC **(d)** 11.08., 06 UTC

(e) 11.08., 12 UTC **(f)** 11.08., 12 UTC

(g) 13.08., 12 UTC **(h)** 13.08., 12 UTC

Figure 5.23: Analyses from 10.08., 12 UTC to 13.08., 12 UTC: Mean (left) and standard deviation (right) OMA and OMB for the assimilated IASI channels. The average for each channel is calculated according to the number of pixels.

5.4.2 Impact results

The impact of assimilating IASI and SYNOP data on the 6-hour forecasts over the European domain is investigated on the low solar (days 03, 04, 09 and 11) and wind (days 09, 10, 11 and 13) power predictability days.

Fig. 5.24 shows the IASI and SYNOP observation impact for the European domain, normalized by observation number, on the 6-hour forecast errors at assimilation times, 00 (Fig. 5.24a), 06 (Fig. 5.24b), 12 (Fig. 5.24c) and 18 (Fig. 5.24d) UTC. On the corresponding assimilation times, the impact of SYNOP and IASI data is plotted for the six days of low wind and solar power predictability. Larger, absolute values on the negative range imply larger improvement of the 6-hour forecast error due to assimilation and larger positive values imply larger forecast degradation. The spatial coverage of IASI data, in combination with the quality control failure (see Section 4.2.1), result in the absence of IASI observation impact seen in Fig. 5.24a and Fig. 5.24b. In Fig. 5.24d, the FSO results showed that the assimilation of IASI and SYNOP on 11.08. at 18 UTC deteriorated the forecast at 00 UTC. This is due to the miss-approximated linear error evolution on 11.08. thus the results are not trustworthy. As described in section Section 5.1 the cause of this might be the number and position of the assimilated observations for the reference run at 00 UTC. In addition, the adjoint model runs failed to give results on 03.08. and 04.08. at 18 UTC. As can be understood the FSO runs at 18 UTC are problematic in contrast to FSO runs at 06 UTC and 12 UTC.

The forecast errors at 06 and 12 UTC, that are found to have the largest RMSE for forecasted (by Germany's TSOs) solar and wind power in the examined case study, come in focus. The assimilation of IASI and SYNOP data at 06 and 12 UTC improved the 6-hour forecast as seen in Fig. 5.24b and Fig. 5.24c. The forecasts benefit more from the assimilation of the data at 06 UTC in comparison to 12 UTC as can be deduced by the range of the negative values on the y-axis. The dramatic error reduction by assimilation of SYNOP data at 06 UTC on 09.08., stands out in Fig. 5.24b. This was also the largest error reduction of the case study period as was found in Chapter 5, Fig. 5.8. The large values of the energy norm in this case, along with the fact of the extreme wind error event on the same day, render FSO as a potential alarm tool that could warn the grid managers of an upcoming extreme error event. Moreover, the ability of the model to accurately predict the weather circulation (see Section 5.3.1) associated with these large error norms can be further investigated in terms of the model itself and the configuration of the observation network.

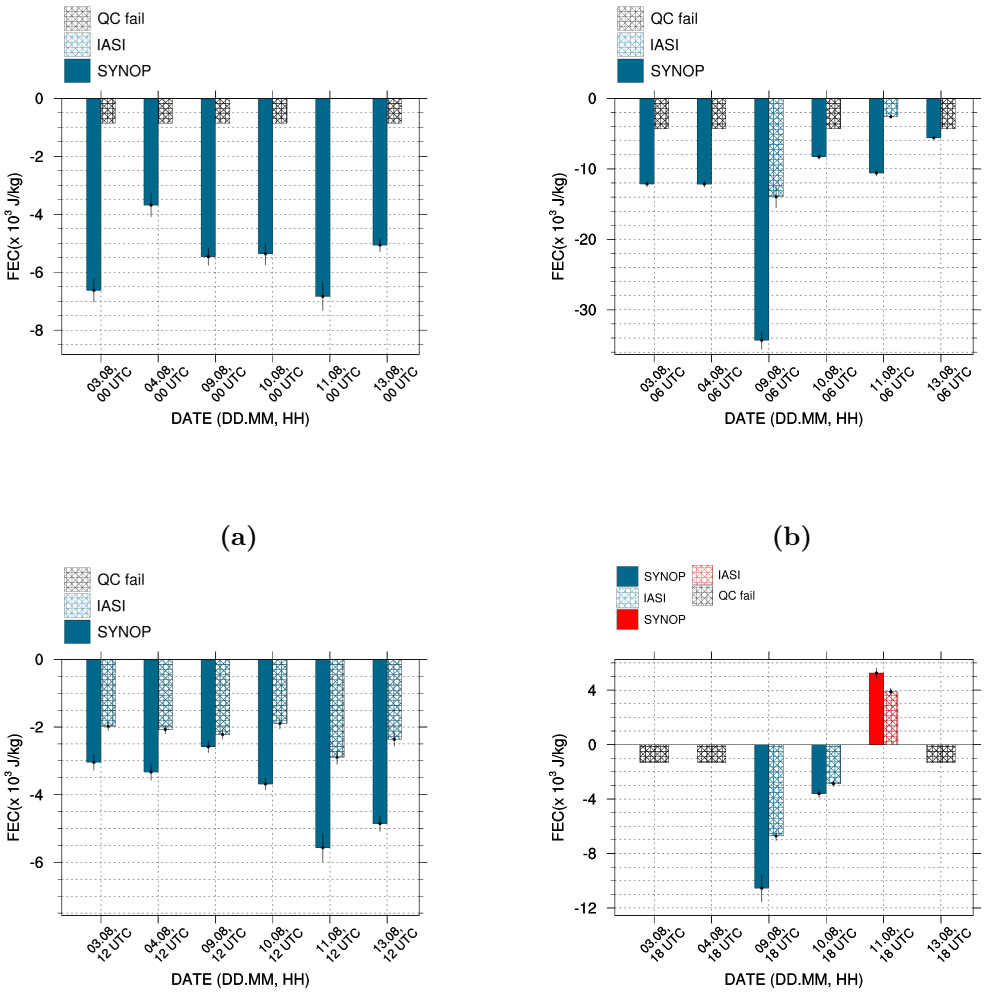
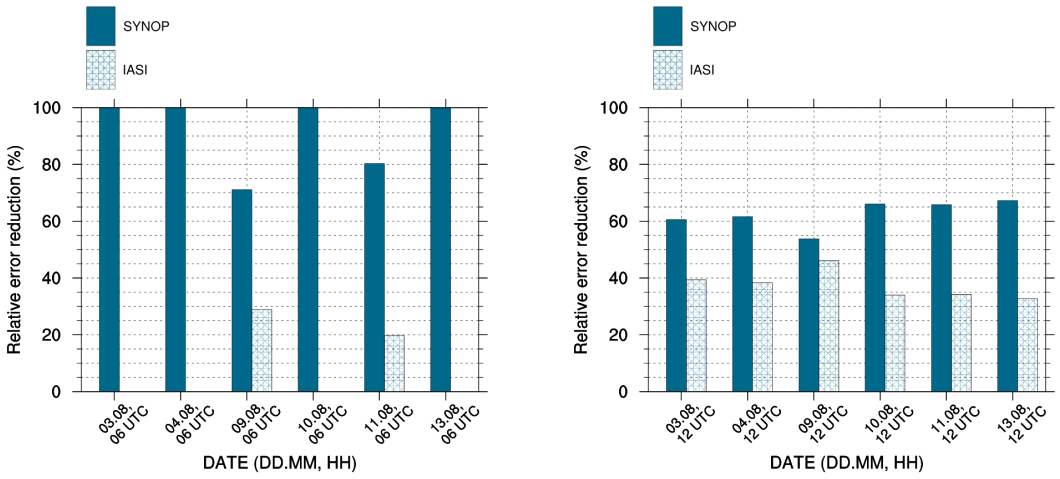


Figure 5.24: Impact of assimilating IASI and SYNOP data on the 6-hour forecasts on 03., 04., 09., 10., 11., and 13.08.2014, for the European domain. The plots aggregate the six days of low solar and wind power predictability per assimilation hour (a) 00, (b) 06, (c) 12 and (d) 18 UTC. Dark blue and red columns indicate the SYNOP data impact and crossed dark blue and red columns described the IASI data impact. The black crossed columns indicate missing impact resluts. The error bars (thin lines) show the mean error defined as $\sigma_{\overline{x}} = \frac{\sigma}{N}$, where σ is the standard deviation and N are the number of observations.

To perceive the magnitude of the contribution to forecast error reduction by IASI and SYNOP data, the relative error reduction (in %) is plotted for assimilation hours at 06 and 12 UTC, in Fig. 5.25. To calculate the percentage, the normalized by observation number, observation impacts are used. At 06 UTC, the conspicuous dominance of SYNOP data accounts for more than 70 % relative error reduction, when both data types are assimilated. On 12 UTC this difference is reduced, with the percentage being ≈ 60 and ≈ 40 % forecast error reduction by SYNOP and IASI data, respectively.

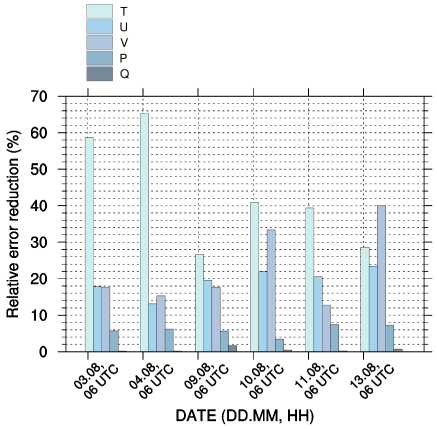


(a) **(b)**

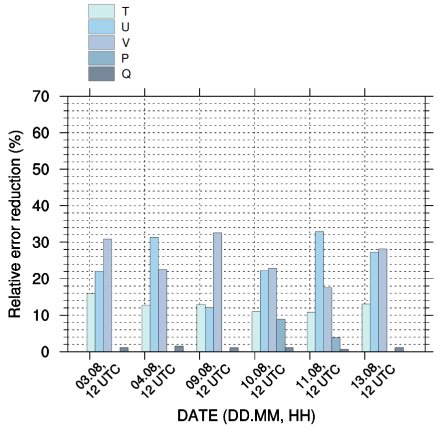
Figure 5.25: Relative error reduction, by assimilation of IASI and SYNOP data at (a) 06 UTC and (b) 12 UTC on the low wind and solar power predictability days.

In Fig. 5.26, the impact of SYNOP and IASI data on the 6-hour forecast error is distributed to their corresponding physical parameters. The relative 6-hour forecast error reduction by assimilation of SYNOP variables and IASI channels is plotted for 06 UTC in Fig. 5.26a,c and for 12 UTC in Fig. 5.26b,d. The y-axis at the same assimilation time for the two observation types, adds up to 100(%) by summing all the SYNOP variables and IASI channels' impact. At assimilation time 06 UTC, the SYNOP temperatures are the most valuable observations (with the exception on 13.08.), followed by the two wind components (Fig. 5.26a). The assimilation of the two wind components dominated the impact results on both hours and it was found that in most cases the assimilation of the meridional wind was more beneficial. This implies larger corrections induced to the model's meridional winds by the assimilation algorithm, which was confirmed by the departure of the observations from the background and the analysis. When IASI is assimilated at 06 UTC, the channels at the water vapor (IASI_Q) and window (IASI_W) region of the infrared spectrum are more important than the temperature channels (IASI_T). Looking at the assimilation hour 12 UTC, the temperature synoptic in-situ observation is not as important as the two wind components. Although IASI is constantly contributing

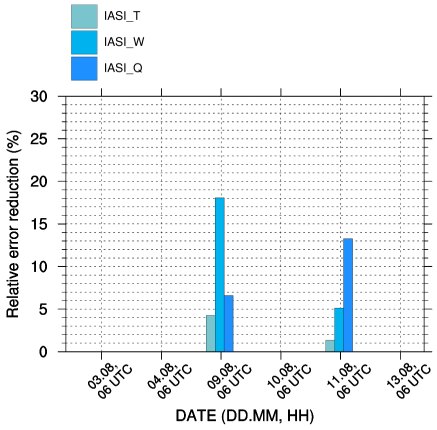
less to the error reduction, it is interesting that the extra 9 water vapor channels, tested for this study, have large contribution to error reduction as seen in Fig. 5.26d.



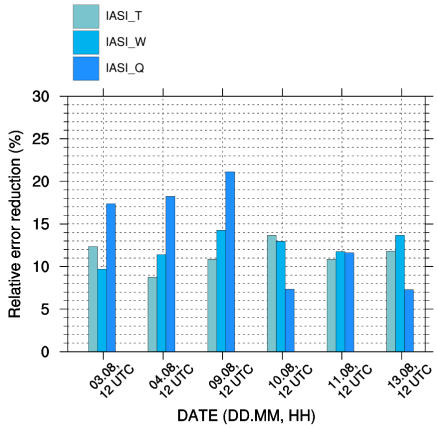
(a) SYNOP assimilation at 06 UTC



(b) SYNOP assimilation at 12 UTC



(c) IASI assimilation at 06 UTC

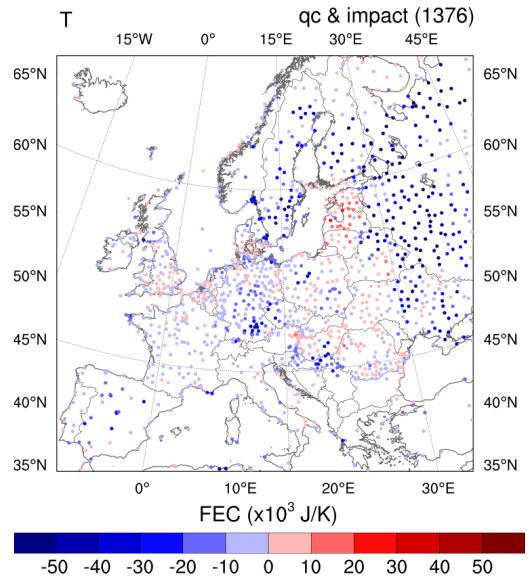


(d) IASI assimilation at 12 UTC

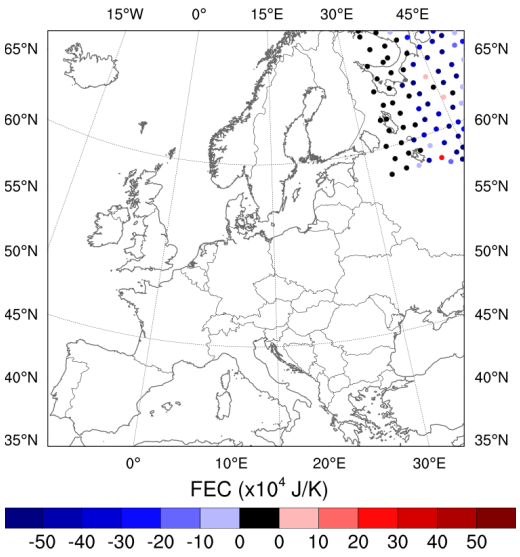
Figure 5.26: Relative error reduction by assimilation of the SYNOP type variables, temperature (T), pressure (P), relative humidity (Q) and the two wind components (U,V) on the 6-hour forecast at (a) 06 UTC and (b) 12 UTC. The relative error reduction by assimilation of IASI is described at (c) 06 UTC and (d) 12 UTC, where $IASI\_T$, $IASI\_W$ and $IASI\_Q$ stand for impact of the channels in the Temperature, Window and Water Vapor regions of the infrared spectrum, respectively.

The forecast error contribution by SYNOP stations providing temperature observations, accounts for ≈ 25 % and ≈ 40 % relative error reduction on the 6-hour forecasts by assimilation on 09.08., 06 UTC and 11.08., 06 UTC, respectively. The position of the most beneficial stations can be located in the maps in Fig. 5.27a and Fig. 5.28a. In the spatial distribution, the observation impact is categorized in bins to create a blue and red contrast, where darker blue represent larger improvement and darker red larger forecast degradation. A symmetric colorbar is chosen. The maximal observation impact out of the six days can be spotted in $36.7°N$ and $28.78°E$ in Fig. 5.27a. It must be noted that negative values smaller than -$50x10^4 J/Kg$ exist. In both plots, the stations on the upper right part of the domain are the most valuable whereas the ones in central Europe even degrade the 6-hour forecasts at assimilation on 09.08.. From the total number of assimilated temperatures (one per station), 26 % and 11% are causing forecast degradation by assimilation on 09.08. and 11.08., 06 UTC, respectively. This degradation is concealed in the previous figures by the larger forecast improvement from assimilation of the remaining stations. The spatial distribution for the impact of SYNOP measured winds on the 6-hour forecast, can be found in Appendix A.2.

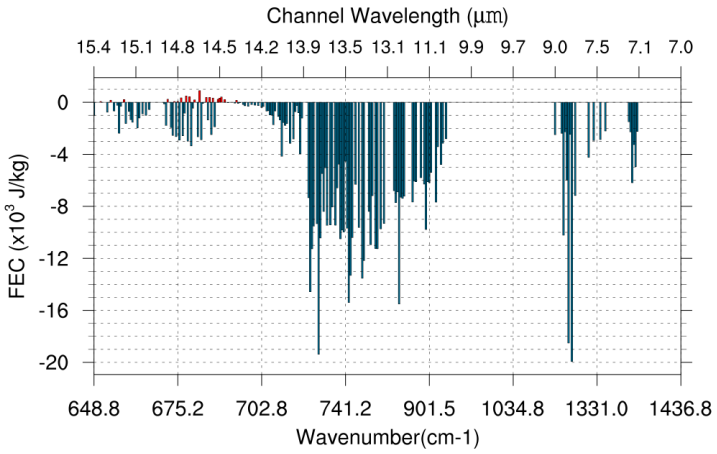
The impact per IASI pixel, as seen in Fig. 5.27b and Fig. 5.28b, is simply the contribution of the assimilated channels to each pixel. Because of this, the colorbars have one order of magnitude larger range. Even with few pixels (in each maximal 158 channels are assimilated), because of the limited spatial coverage at 06 UTC, the forecasts benefit from IASI assimilation. The ranking list of the most beneficial IASI channels, normalized by pixel number, can be seen in Fig. 5.27c and Fig. 5.28c. Despite more temperature channels ($15.4 - 13.5\mu m$) being assimilated, the channel group is not as beneficial for assimilation as the other two channel groups. This is also evident in Fig. 5.26c, with more than 15 % relative error reduction being attributed to channels in the window region of the infrared spectrum on 09.08., 06 UTC and 11 % to the water vapor region on 11.08., 06 UTC. The error reduction by the assimilation of temperature channels at 06 UTC, on the examined days, is less than 5 %. From the 102 temperature channels assimilated on 09.08. at 06 UTC, 20 channels were mildly (compared to the forecast improvement range) degrading the forecast. The assimilation of 12 channels on 11.08. at 06 UTC was found to be deteriorating the forecast, with largest forecast errors imposed by three water vapor channels. This is in contrast to the good results obtained by assimilating the water vapor channels.



(a) SYNOP, 09.08.

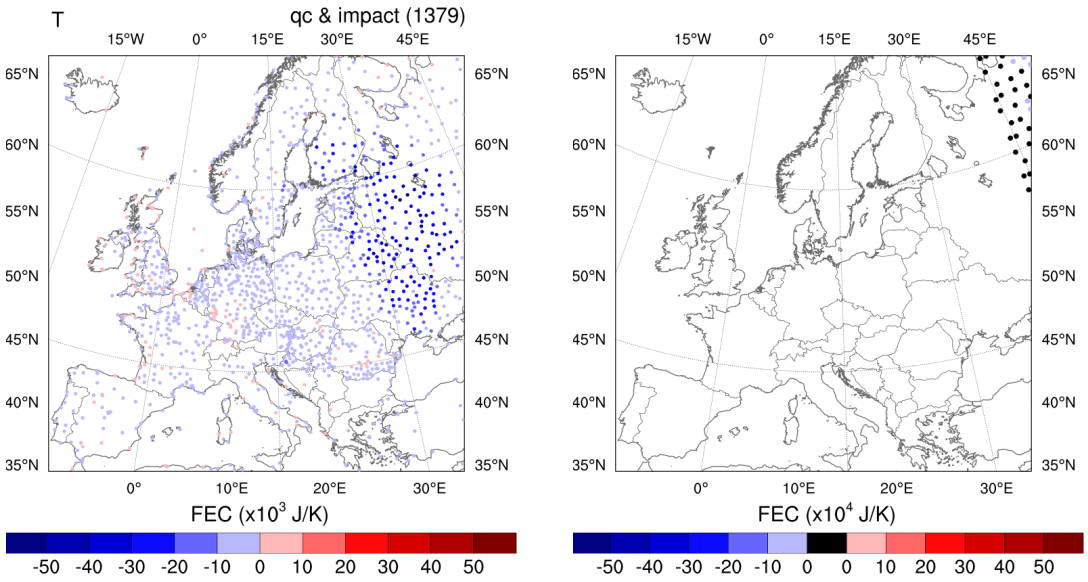


(b) IASI, 09.08.



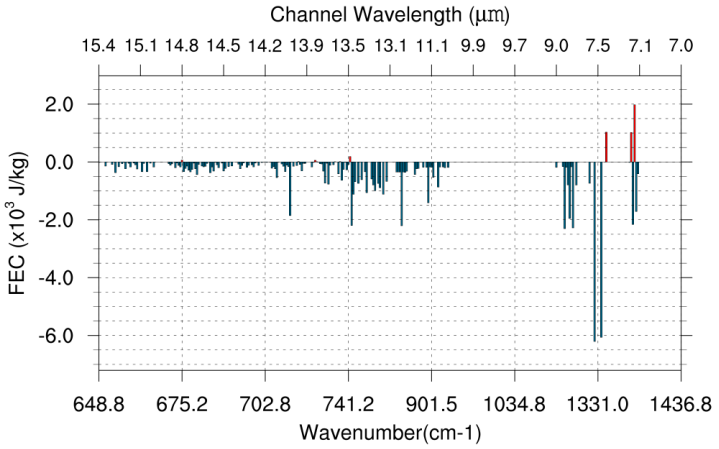
(c) IASI, 09.08.

Figure 5.27: Impact of (a) SYNOP temperature (b) IASI pixels (c) IASI channels assimilation on 09.08. 06 UTC on the 6-hour forecast. The impact per channel is normalized by number of pixels. Black pixels describe quality control failure. The colorbars are generated by taking the minimum and maximum impact for the six examined days with an interval of ten. The left and right edges indicate values smaller than -50 ($\mathbf{x}10^3$ for SYNOP and $\mathbf{x}10^4$ for IASI) and greater than 50 ($\mathbf{x}10^3 J/Kg$ for SYNOP and $\mathbf{x}10^4 J/Kg$ for IASI), respectively.



(a) SYNOP, 11.08.

(b) IASI, 11.08.

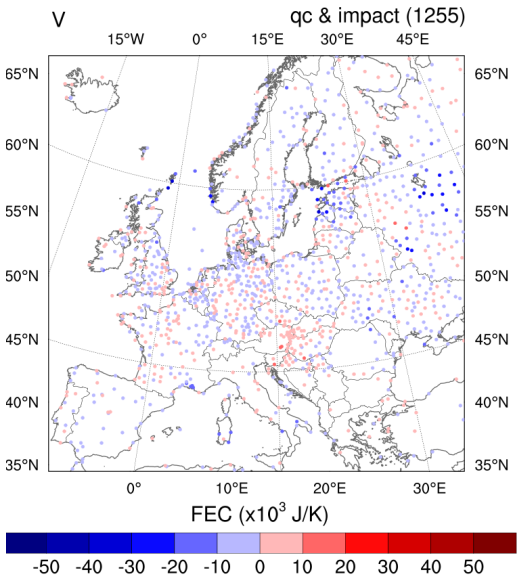


(c) IASI, 11.08.

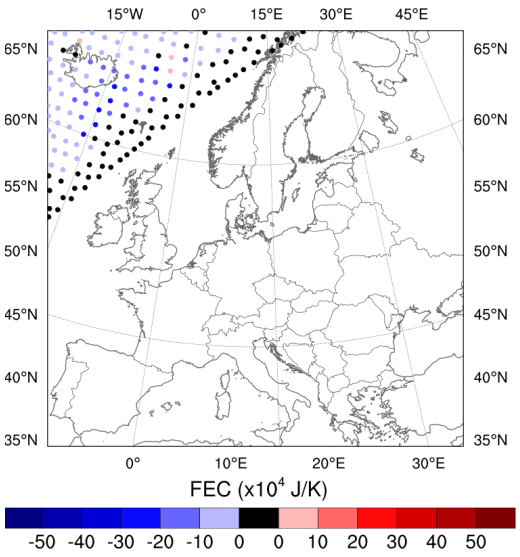
Figure 5.28: Impact of (a) SYNOP temperature (b) IASI pixels (c) IASI channels assimilation on 11.08. 06 UTC on the 6-hour forecast. The impact per channel is normalized by number of pixels. Black pixels describe quality control failure.

In the following Section 5.4.2 to Fig. 5.34, the same plot types are utilized to describe the FSO results on assimilation at 12 UTC on days of low wind and solar power predictability. On all days, ≈ 50 % of the relative error reduction is attributed to the assimilation of SYNOP wind components (Fig. 5.26b). For each date, the observation impact of the largest contributing SYNOP variable is plotted on the map. The mild forecast degradation and forecast improvement per station is a common feature in all SYNOP stations, with individual stations compensating with large absolute negative values. The percentage of stations declining the forecast varies from 24 % on 04.08. to 41 % on 10.08.. A cluster of stations located in northern-eastern Germany, is found to constantly improve the forecast.

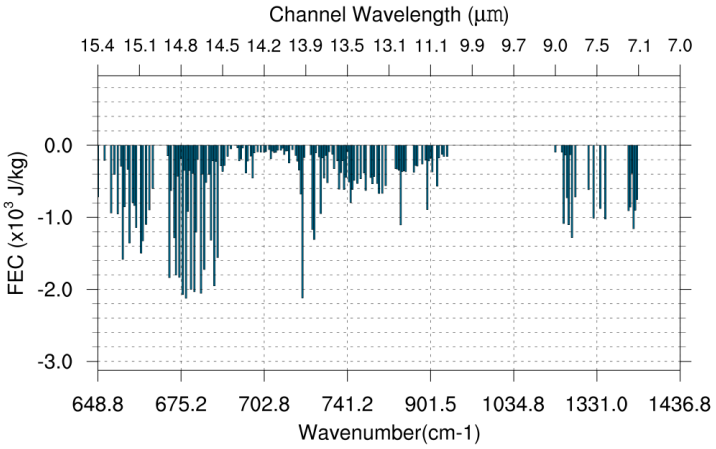
Taking a closer look at the IASI impact per pixel, it is found that less than 2.5 % of the pixels have positive values on all days and that the prevailing impact range is between 0 and -10 $x10^4 J/Kg$ with the exception on 11. and 13.08. On those days, most of the pixels are binned between the larger forecast error reduction range, -20 to -10 $x10^4 J/Kg$. When examining the contribution per group of channels, it is found that the IASI_Q have a larger contribution to the relative error reduction on days 03., 04., and 09.08., which reaches up to 20 % on 09.08. (Fig. 5.26d). This can be explained, by the large forecast improvement imposed by them, relative to their number (9 channels). The individual channel contribution of the IASI_T is well noticed in all plots of channel impact, normalized by pixel number (Fig. 5.29c to Fig. 5.34c). Especially the ones in the range 15.4 -14.5μm are consistently the most valuable in the temperature channel range. On the other hand, insignificant forecast improvement is found for all days by assimilation of channels of the temperature group between 14.5 - 14.0 μm. On 09.08., 10 channels in the temperature region had small positive values, mostly from that range.



(a) SYNOP, 03.08.

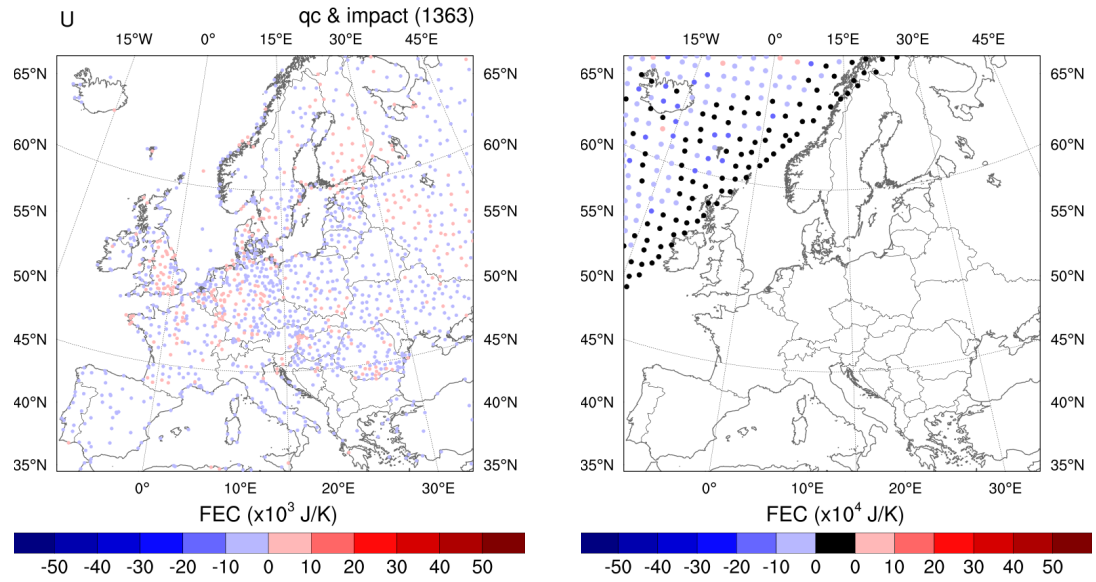


(b) IASI, 03.08.



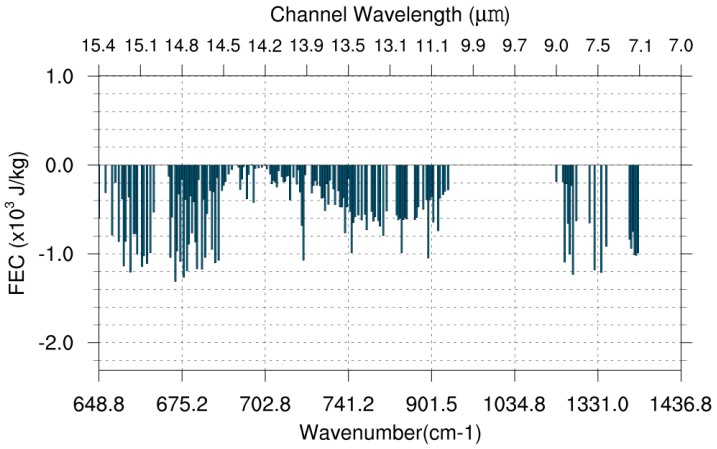
(c) IASI, 03.08.

Figure 5.29: Impact of (a) SYNOP temperature (b) IASI pixels (c) IASI channels assimilation on 03.08., 12 UTC on the 6-hour forecast. The impact per channel is normalized by number of pixels. Black pixels describe quality control failure.



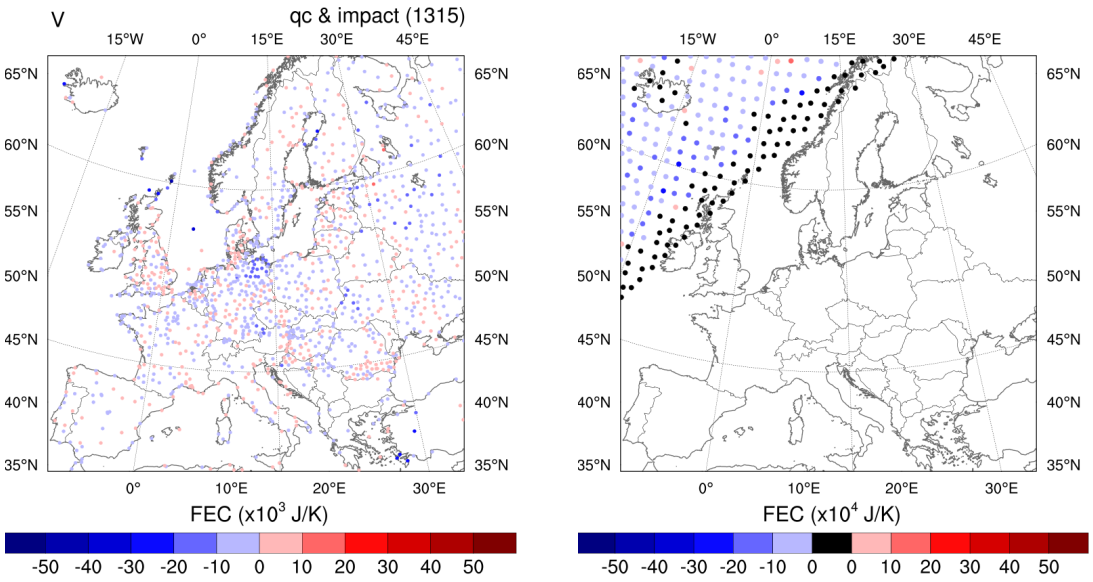
(a) SYNOP, 04.08.

(b) IASI, 04.08.



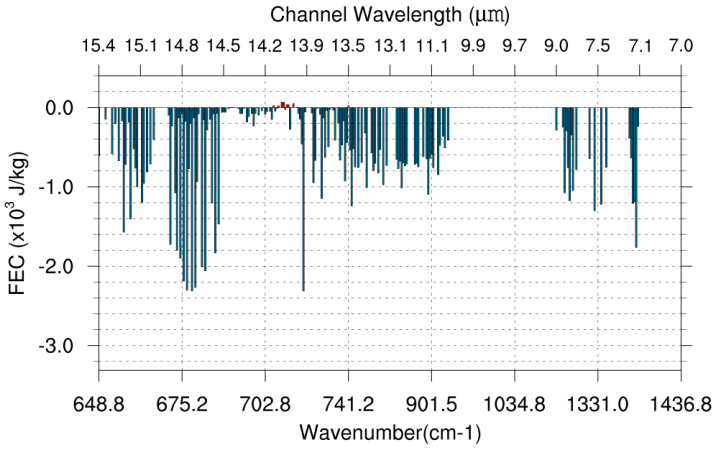
(c) IASI, 04.08.

Figure 5.30: Impact of (a) SYNOP temperature (b) IASI pixels (c) IASI channels assimilation on 04.08., 12 UTC on the 6-hour forecast. The impact per channel is normalized by number of pixels. Black pixels describe quality control failure.



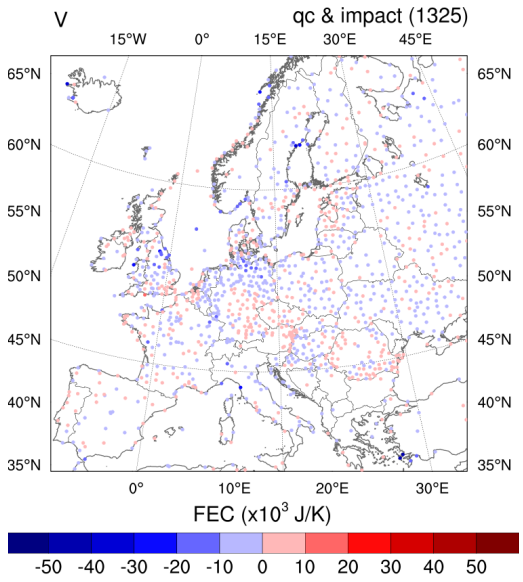
(a) SYNOP, 09.08.

(b) IASI, 09.08.

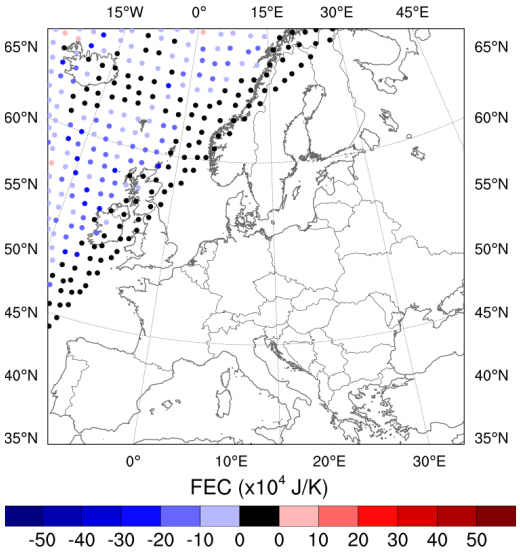


(c) IASI, 09.08.

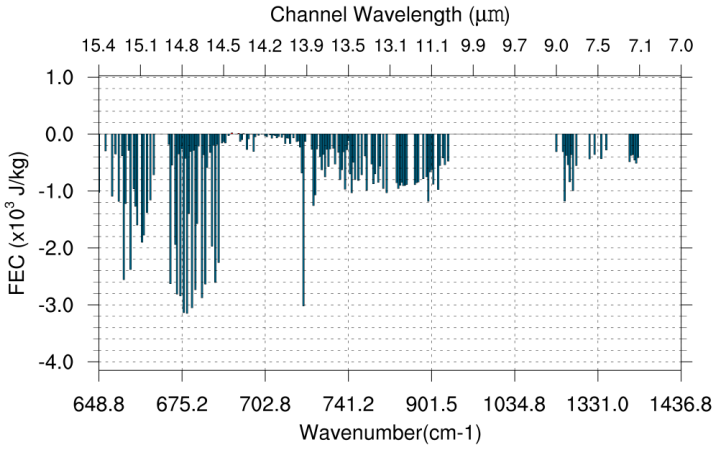
Figure 5.31: Impact of (a) SYNOP temperature (b) IASI pixels (c) IASI channels assimilation on 09.08., 12 UTC on the 6-hour forecast. The impact per channel is normalized by number of pixels. Black pixels describe quality control failure.



(a) SYNOP, 10.08.

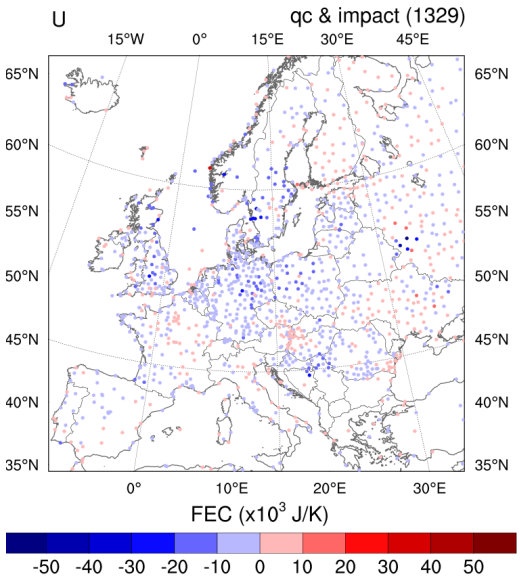


(b) IASI, 10.08.

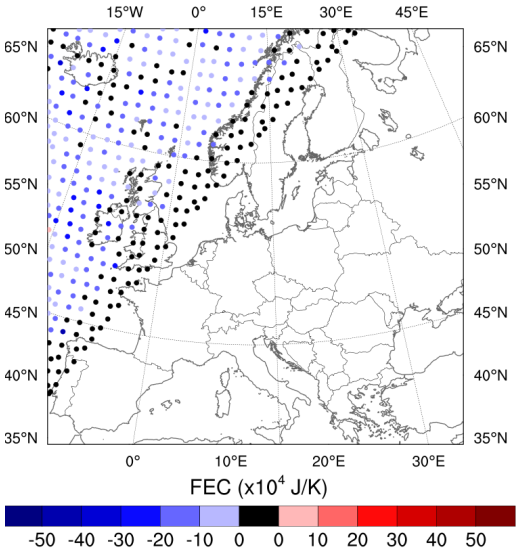


(c) IASI, 10.08.

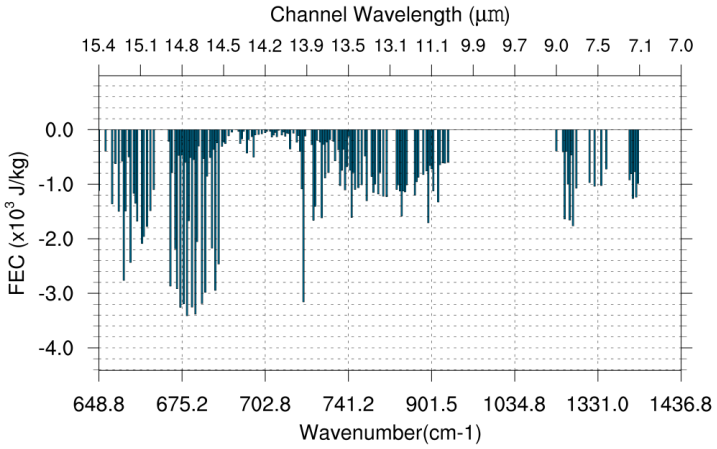
Figure 5.32: Impact of (a) SYNOP temperature (b) IASI pixels (c) IASI channels assimilation on 10.08., 12 UTC on the 6-hour forecast. The impact per channel is normalized by number of pixels. Black pixels describe quality control failure.



(a) SYNOP, 11.08.

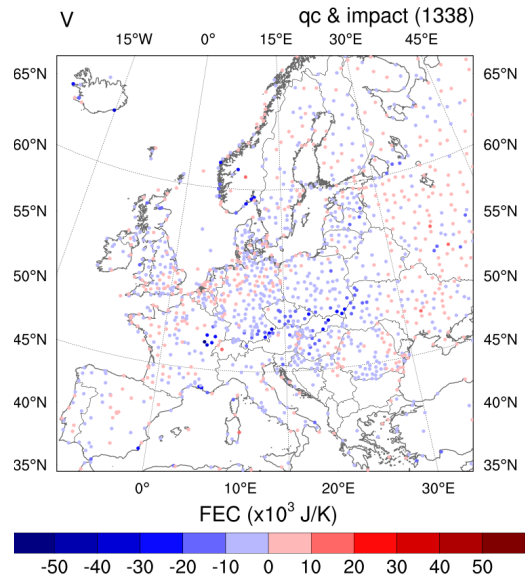


(b) IASI, 11.08.

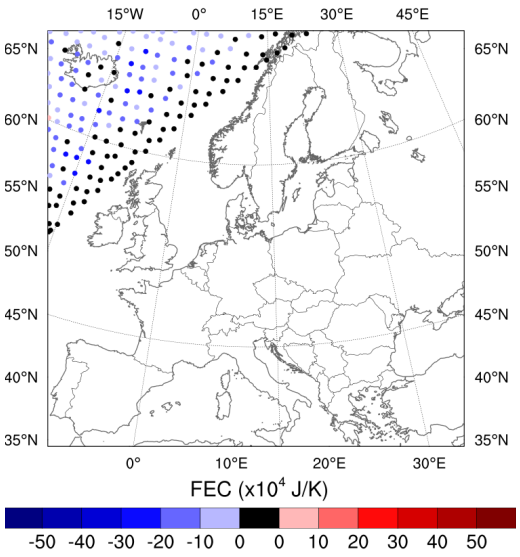


(c) IASI, 11.08.

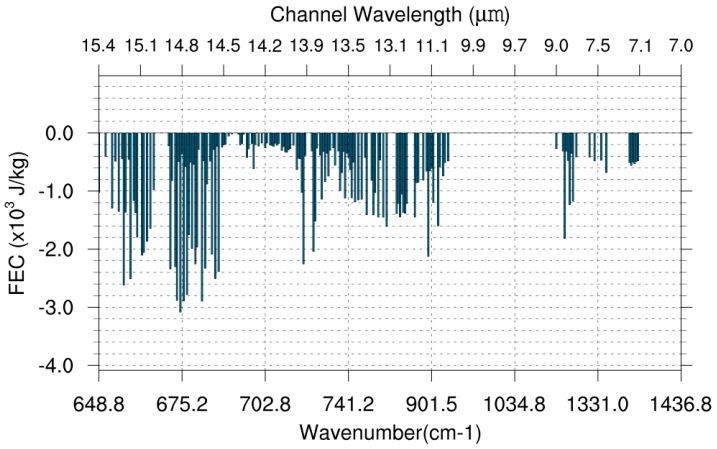
Figure 5.33: Impact of (a) SYNOP temperature (b) IASI pixels (c) IASI channels assimilation on 11.08., 12 UTC on the 6-hour forecast. The impact per channel is normalized by number of pixels. Black pixels describe quality control failure.



(a) SYNOP, 13.08.



(b) IASI, 13.08.



(c) IASI, 13.08.

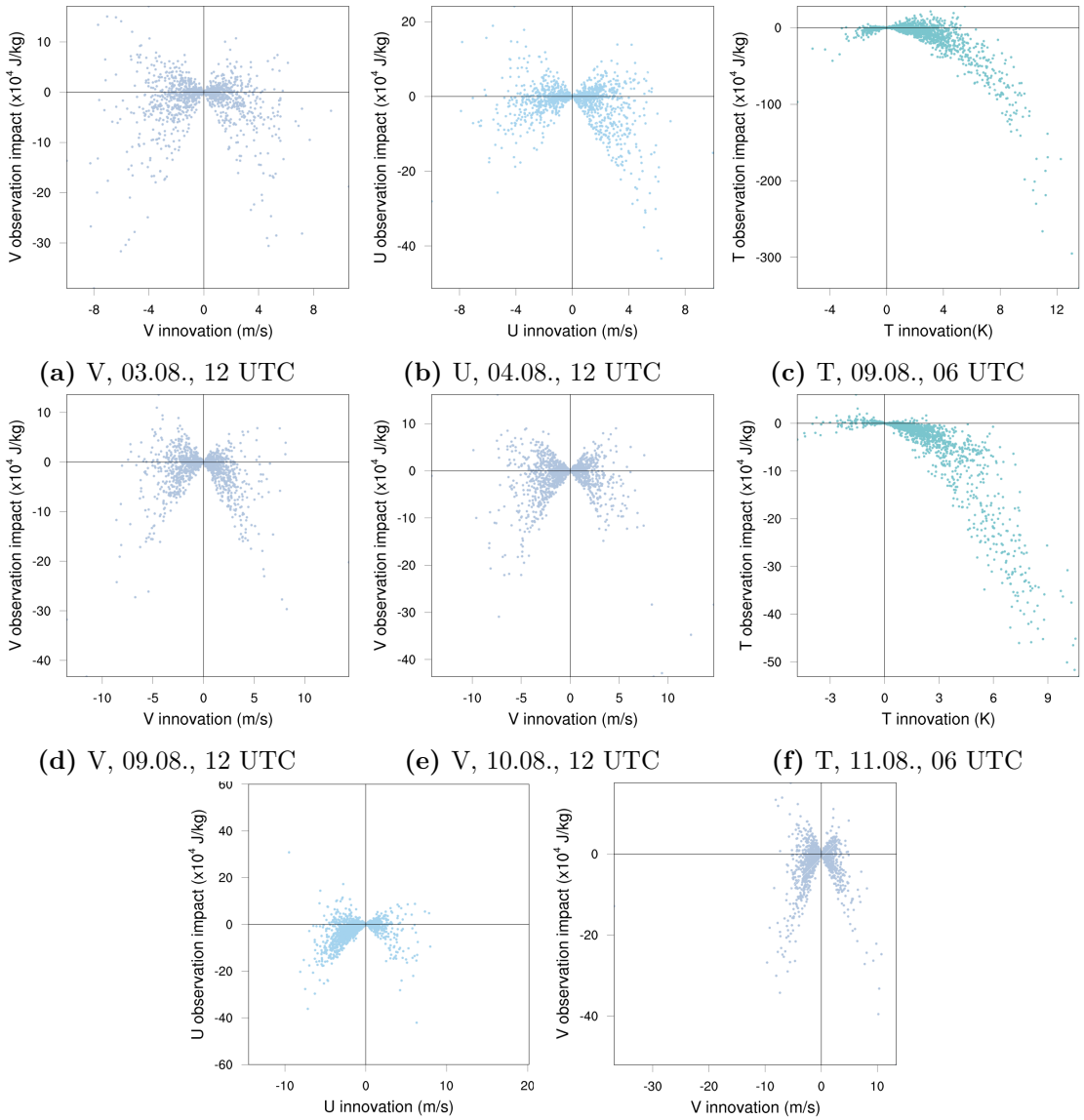
Figure 5.34: Impact of (a) SYNOP temperature (b) IASI pixels (c) IASI channels assimilation on 13.08., 12 UTC on the 6-hour forecast. The impact per channel is normalized by number of pixels. Black pixels describe quality control failure.

5.4.3 Examination of observation impact

The scale at which each variable contributes to the observation impact can be understood by the relationship of the observation impact and the innovation vector $\mathbf{d}$ (see Eq. (3.2)). It is clear from the definition of the observation impact (or FEC) in Eq. (2.19), that the greater the distance of the observation from the background information, the greater the observation impact. The positive and negative sign of the impact reveals if the 6-hour forecast originating from the analysis is farther or closer to the reference state than the forecast without assimilation, respectively.

This relationship is examined through the following scatter plots of observation impact and innovation vector. The focus lies on the six low predictability days and the most important contributing SYNOP variables (Fig. 5.35) and the three IASI channels groups (see Fig. 5.36). In the scatter plots, observations below the y-axis zero reference line are the ones improving the 6-hour forecast and above the axis are the ones deteriorating the forecast. Observations left of the x-axis zero reference line have negative innovation, meaning the model over-predicts the value of the observation.

At a first look at Fig. 5.35, the symmetry about the y-axis is noticed for the wind components scatter plots (Fig. 5.35a, b, e and g). This means that, as the absolute innovation increases, so does the absolute observation impact. In the temperature scatter plots in Fig. 5.35c and Fig. 5.35f, a tendency of larger forecast improvement by large model under-predicted observations is noticed in the lower right parts of the graphs. In Fig. 5.35c, larger degradation due to temperature assimilation is noticed, compared to Fig. 5.35f. In this case the model is under-predicting the observations (upper right part on the graph). The opposite is true for Fig. 5.35f, as deduced by comparing the upper right and left parts of the graph. In all figures, the y-axis range has larger negative than positive range, which is expected as these are the variables with the best performance.



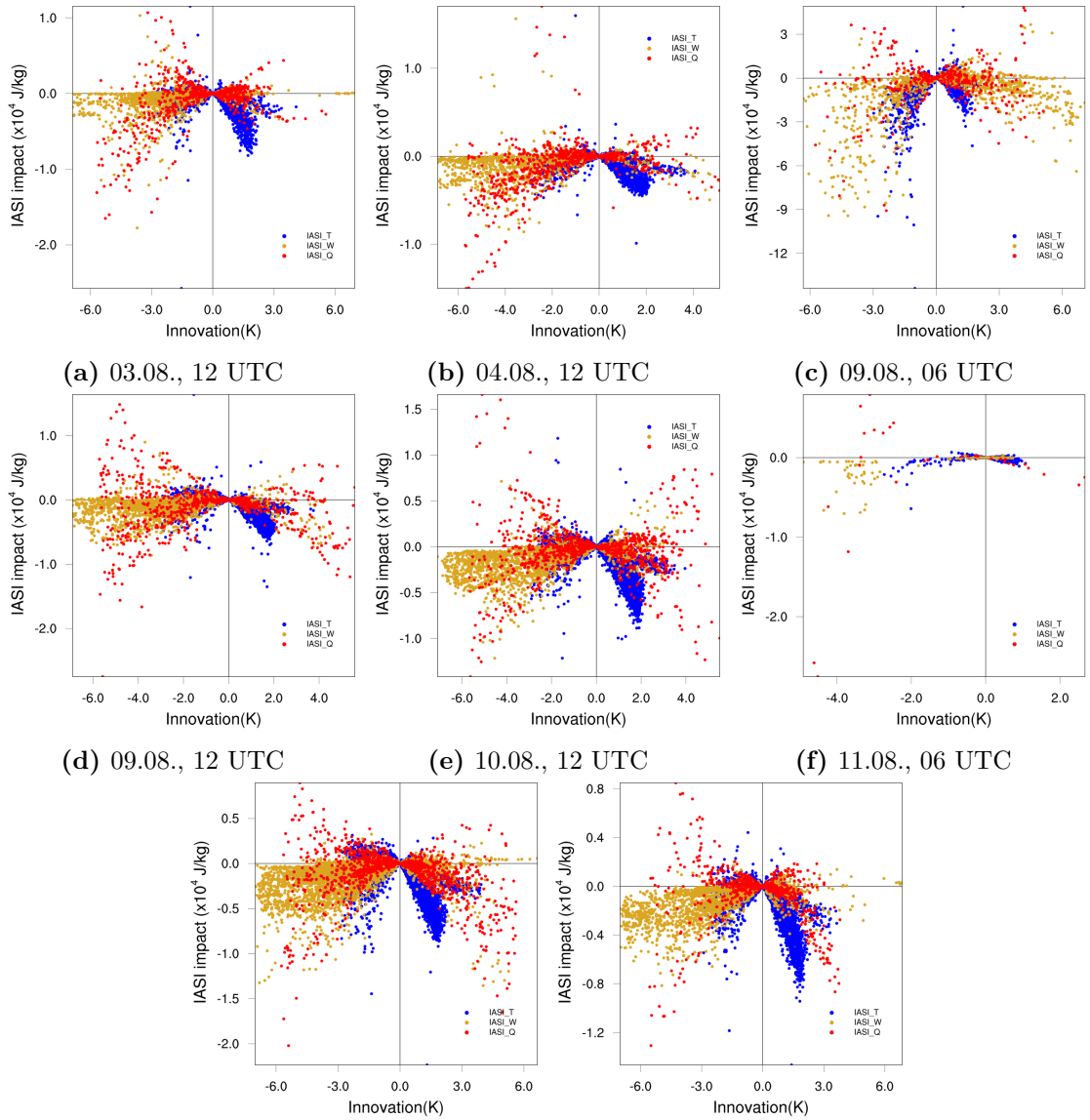
(a) V, 03.08., 12 UTC **(b)** U, 04.08., 12 UTC **(c)** T, 09.08., 06 UTC

(d) V, 09.08., 12 UTC **(e)** V, 10.08., 12 UTC **(f)** T, 11.08., 06 UTC

(g) U, 11.08., 12 UTC **(h)** V, 13.08., 12 UTC

Figure 5.35: Scatter plots of observation impact of SYNOP variables and their corresponding innovation vector. The most valuable SYNOP variable, for the 6-hour forecast on each low predictability day, is shown.

The scatter plots of IASI channels and innovation vector are plotted in Fig. 5.36. The number of observations in the plots, depends on the number of assimilated pixels. Each channel group, IASI_T, IASI_W and IASI_Q, is plotted blue, yellow and red, respectively. The smaller impact values on the y-axis compared to SYNOP scatter plots is apparent. This confirms the qualitative superiority of synoptic measurements in the forecast improvement. The three channels groups are plotted together on each day in order to compare with the relative error reduction by each group shown in Fig. 5.25b. For example the larger contribution of the IASI_Q channels in the 6-hour forecast error reduction by assimilation on 03 and 04.08. at 12 UTC can be seen in Fig. 5.36a and Fig. 5.36b. Specifically, the large error reduction can be seen in the lower left part on the graph, where the model over-predicts the observations up to 6K. On Fig. 5.36f, few pixels are assimilated and the forecast degeneration by the IASI_Q is seen in the upper left part of the graph, were the model over-predicts the observations by $\approx$ 3K. The specific channels that caused the degradation are identified in Fig. 5.28c. The IASI_T channels, have smaller innovation vectors compared to the other groups. This can be understood, as the radiative transfer model is well prepared for this range of infrared spectrum and does not involve the non-linearities introduced by the water vapor channels (as discussed in Section 4.1.1). The IASI_W, persistently cause the largest error degradation when the observations are over-predicted up to 6K (lower left in all figures).



(a) 03.08., 12 UTC **(b)** 04.08., 12 UTC **(c)** 09.08., 06 UTC

(d) 09.08., 12 UTC **(e)** 10.08., 12 UTC **(f)** 11.08., 06 UTC

(g) 11.08., 12 UTC **(h)** 13.08., 12 UTC

Figure 5.36: Scatter plots of the IASI channels observation impacts and their corresponding innovation vector. All three IASI groups (IASI_T, IASI_W and IASI_Q) are plotted in each low predictability day with blue, yellow and red, respectively. The number of observations depends on the number of pixels assimilated on each day.

The results obtained demonstrate the relation of the assimilation system and the impact of observations on the short-term forecasts. Tendencies of over/under prediction by the model can be identified and can be used for the benefits of accurate predictions. This can be particularly important for extremely poor solar and wind power forecasts. For example, the impact results showed that the degradation of the 6-hour forecast due to assimilation of stations measuring temperature on 09.08. at 06 UTC could be located in central Europe. The majority of those temperature observations was over-predicted by the assimilation model. The assimilation of the synoptic winds, was found to be the most beneficial component of the examined observation network. On the days of low wind predictability, the assimilation model was constantly over-predicting the winds. The assimilation of SYNOP humidity was found to be negligible. This was not the case for the IASI water vapor channels, peaking in the mid-upper troposphere. With the exception of assimilation on 11.08. at 06 UTC, the nine water vapor channels were found to add substantial value to forecast improvements. The largest improvements by assimilation of the water channels coincided with the days of low solar power predictability (days 03., 04. and 09.08.).

Of course, the results apply for this observation network and model set-up. The quantity and quality of observation types, the assimilation model set-up can alter the results. Nonetheless, if the former are chosen then FSO can be set to monitor the selected observation network and hint for preferences of the assimilation system for over or under predicting the observations. This can be related then to weather circulation patterns and the model's ability to accurate predict it. The alarm mode of FSO for extreme power error events is shown here as a computationally cheap monitoring tool, that based on the selected observation network and one can define an energy norm above which the grid managers can be warned.

Chapter 6

Conclusions

Data assimilation can serve energy-meteorology by improving forecasts in support of resilient power grid management and realistic price fixing for electrical power trading. On its part, data assimilation can be improved by evaluating the observation network configuration. This is demonstrated in this work, where the FSO algorithm is applied, on low solar and wind power predictability days and the information value of IASI and SYNOP data is ranked for short term forecasts. A satellite sensor versus ground-based configuration scheme was chosen and the 6-hour weather forecast error reduction was calculated in a competing mode, for the European domain, and the case study period 01.-15.08.2014. This competing mode, allows the creation of a ranking list of the most beneficial observation types and measured variables, dependent on weather situations. Most importantly, the impact of individual observations or observation types can be isolated to particular stations or channels. Taking advantage of this feature, the observation network configuration can be evaluated for days of low wind and solar power predictability along with the associated atmospheric circulation.

The length of two weeks for the experiment was selected in order to evaluate the results with OSEs and the specific dates were selected to include the extreme wind error event on 09.08.2014. The four most erroneously forecasted days were identified for solar (days 03, 04, 09 and 11) and wind power (days 09, 10, 11 and 13) according to data of Germany's TSOs. The largest RMSE for both solar and wind power was found on 09.08.2014. The RMSE, for this day, was found to reach 4.03 GW and 1.37 GW for wind and solar power, respectively. This translates into 34% and 16% under-prediction of the actual energy production by the forecasts. The dominant weather evolution on those days, were strong baroclinic developments with upper level troughs.

The aggregated results for the case study period, revealed that the assimilation of these two observation types reduces on average the 6-hour forecast error. IASI was found to contribute twice as much on the average error reduction comparing to SYNOP. The results were supported by the conducted OSEs, which qualitatively showed a reduction of the averaged forecast error (not normalized) when both observation types were assimilated and increased forecast errors on absence of IASI data, on all assimilation hours. The relation of observation number and impact

was clear, with the bulk IASI data and especially the temperature channels being identified as more beneficial. Upon normalisation by the number of observations the impact results revealed the larger contribution to forecast error reduction by SYNOP. The normalized values showed that 70% of the relative averaged error reduction is attributed to SYNOP data, from which more than 40% is due to assimilation of the wind components. The introduction of the nine water vapor channels, was found to be double as beneficial for the averaged forecast error reduction compared to the assimilation of the temperature channels, when examining the normalized by observation number impact.

Shifting the focus to the low predictability days, the most valuable SYNOP variables and IASI group channels were identified. Larger error reduction was induced by assimilation of the two observation types at 06 UTC, compared to 12 UTC. The assimilation of IASI and SYNOP on 09.08. at 06 UTC had the largest 6-hour forecast error reduction of the examined period. This coincided with the day of the extreme error of wind power forecast. The assimilation of the two synoptic wind components dominated the impact results on both hours and it was found that in most cases the assimilation of the meridional wind was more beneficial. For the majority of the examined days, the model was on average over-predicting the SYNOP wind observations. The water vapor channels were found to rank among the most beneficial observations on the low solar power predictability days, reaching 20% relative error reduction on 09.08. at 12 UTC.

Next, the impact was spatially investigated, revealing locations of SYNOP stations and IASI pixels that improve or degrade the forecast. The most prominent contrast of synoptic stations that improved or degraded the forecast was found for the synoptic temperature assimilation on 09.08., 06 UTC. A cluster of stations (26% of the total number of stations) positioned in central Europe was degrading the 6-hour forecast. This was concealed in the total temperature impact by the larger forecast improvement from assimilation of the remaining stations. It is unclear whether this degradation was due to poor observations, false forecasts or even untrustworthy reference state. Unfortunately FSO can not distinct which of the above cases caused the forecast deterioration but it can alarm that this set-up, with this assimilation and forecast model and these observations is degrading the 6-hour weather forecast. The mild forecast degradation and forecast improvement per station was a common feature for the majority of SYNOP stations, with individual stations compensating with large absolute negative values. The percentage of stations, declining the forecast (calculated for the largest contributing SYNOP variable to the forecast error) varied from 24% on 04.08. to 41 % on 10.08.. A cluster of stations located in North-East Germany, was found to constantly improve the forecast on all examined days.

Even with few pixels, the forecasts benefit from IASI assimilation. It was found that less than 2.5 % of the pixels had positive (= degradation) values on all examined days. The impact per individual channel, normalized by number of pixels, disclosed particular channels in the range of $15.4 - 14.5\mu m$ (part of the IASI_T channel group) that had the largest contribution to forecast error reduction. On the other hand, channels in the range of $14.5 - 14.0\mu m$ (also part of the IASI_T channel group) had an insignificant impact. These channels were constantly under-predicted

by the assimilation algorithm and had the smallest corrections (Obervation-Minus-Background). The range of the impact values was linked to the distance of the observations from the model background, where larger distances imposed larger observation impact values.

Though some spatial patterns were identified for SYNOP stations, no conclusion can be safely drawn for the optimal spatial configuration of the stations on the low predictability days. This is especially hindered by the observation impact range. Being in a competing mode error norm, where larger absolute values indicate larger impact, it is difficult to set a limit or a range for which the impact is considered big enough. What can be done is to define a threshold according to impact results, with the same model and observation configuration, from extended experiments longer than two weeks. This can be particular useful for extreme error events. It can be used, to train the model with the best observation network possible to be ready for these occasions. Of course more observation types, potentially everything that can be assimilated should be included. This should also increase the accuracy of the algorithm, that was found to be problematic when IASI was absent.

The under-prediction of the solar power on the poorest predicted days for solar energy was accompanied with scattered clouds over Germany. The poor ability of the NWPs to simulate the cloud fields, the non-linearities introduced in the assimilation process and the cloud-cleared IASI data deter the effort towards accurate cloud prediction. Despite this, the assimilation of nine water vapor channels, whose weight functions peak in the upper troposphere, were found to be among the most beneficial observation types. The impact of the water vapor channels, normalized by observation number, on forecast error reduction was larger compared to the other two channel groups on all low solar power predictability days. This can lead the way for testing the assimilation of the IRS onboard the future MTG satellite. The humidity, measured by the synoptic stations had negligible impact on the forecast error. The amplitude by which these corrected by the assimilation fields can lead to more accurate cloud representation by the model merits a separate study.

The potential of this algorithm to serve energy meteorology is clear. Despite the limitations that come along, it can be a useful asset for monitoring the observation network. Going one step further, the discrepancy measured in the forecast error between forecast and reference state could be transferred from meteorological metrics to actual power values. This of course would demand a module that relates the forecasted 100 m height wind and solar irradiation with power output.

Appendix A

Appendix

A.1 FSO configuration at 18 UTC

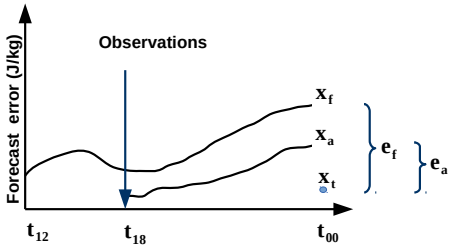
Taking a closer look at the experiment configuration for assimilation at 18 UTC (Fig. A.1a), the non-linear forecast error at 00 UTC originating from the assimilation at 18 UTC would be

$$\mathbf{e_a} = (\mathbf{x_a} - \mathbf{x_t})^{\mathbf{T}}\mathbf{C}(\mathbf{x_a} - \mathbf{x_t}). \tag{A.1}$$

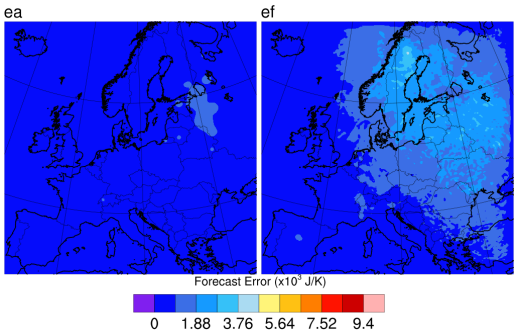
Likewise, the error at 00 UTC from the forecast originating 6 hours before assimilation and used as background in the assimilation would be

$$\mathbf{e_f} = (\mathbf{x_f} - \mathbf{x_t})^{\mathbf{T}}\mathbf{C}(\mathbf{x_f} - \mathbf{x_t}). \tag{A.2}$$

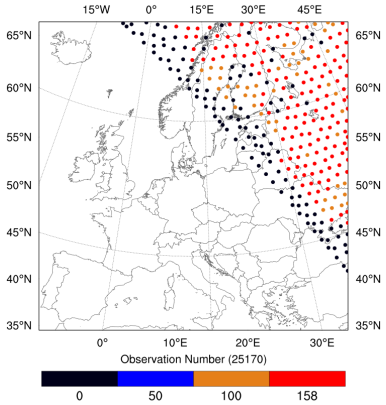
The time average for assimilation runs at 18 UTC and forecast errors at 00 UTC and vertically integrated forecast errors, $\mathbf{e_a}$ and $\mathbf{e_f}$ are plotted in Fig. A.1b. The forecast errors are namely in the upper right corner of the domain for $\mathbf{e_f}$, as this area was not affected by the assimilation of plentiful IASI data at 18 UTC (Fig. A.1c). On the contrary, the correction on the top right corner is evident for $\mathbf{e_a}$. However, because the reference analysis $\mathbf{x_t}$ contains little information from IASI data (Fig. A.1d) the quadratic difference in Eq. (A.1) has relatively small values.



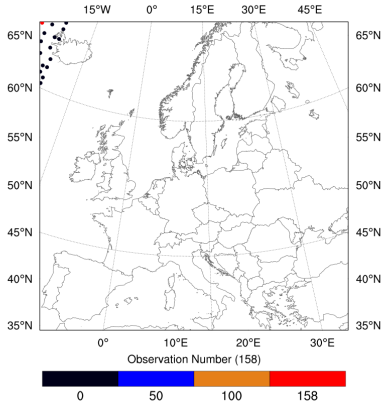
(a) Experiment configuration at 18 UTC



(b) Average forecast error at 00 UTC



(c) IASI coverage at 18 UTC



(d) IASI coverage at 00 UTC

Figure A.1: (a) Experiment configuration for assimilation at 18 UTC. (b) Timely averaged non-linear forecast errors at 00 UTC from analyses at all days at 18 UTC ($\mathbf{e_a}$) and from simple forecasts ($\mathbf{e_f}$). The color-scale has the range of the minimal and maximal forecast errors found for all examined analyses times in order to compare the magnitude of the sensitivity of the adjoint procedure to the error range. (c) IASI spatial coverage at assimilation time 18 UTC and (d) at forecast time at 00 UTC.

This is not the case for the configuration at assimilation at 12 UTC, where the quadratic errors $\mathbf{e_a}$ and $\mathbf{e_f}$ differ, and the corrections by the assimilation are visible.

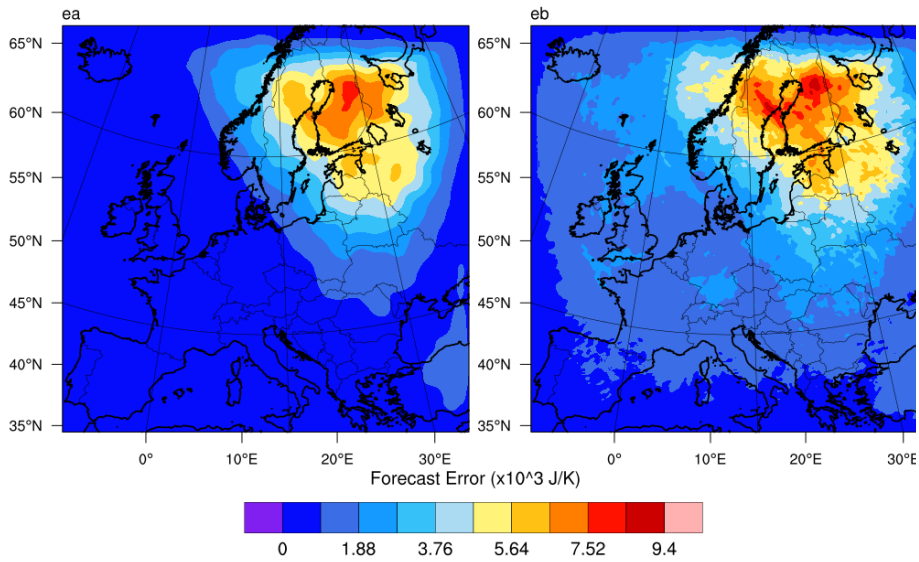
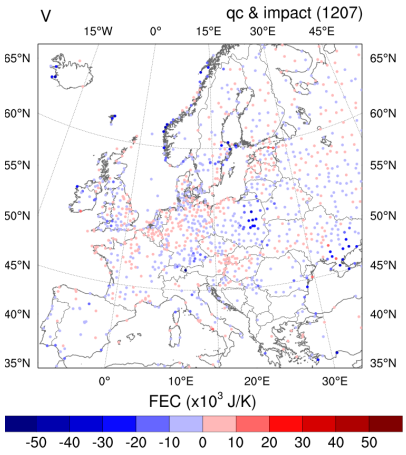


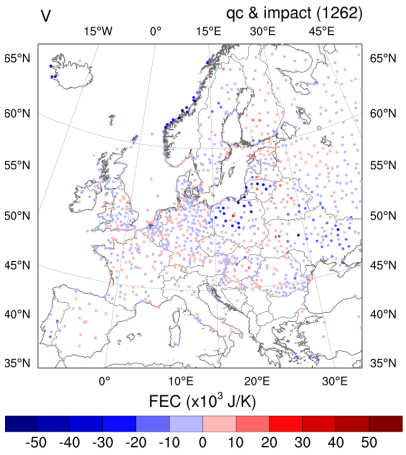
Figure A.2: Timely averaged non-linear forecast errors at 18 UTC and assimilation at 12 UTC.

A.2 Spatial distribution of observation impact for SYNOP wind measurements

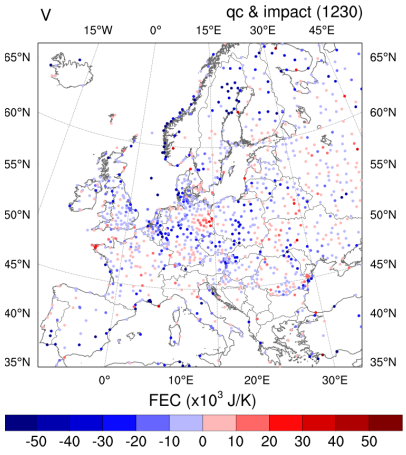
As the wind measurements were the most beneficial for the short-term forecast error reduction, all spatial distribution plots of wind impact on poor predictability cases were created.



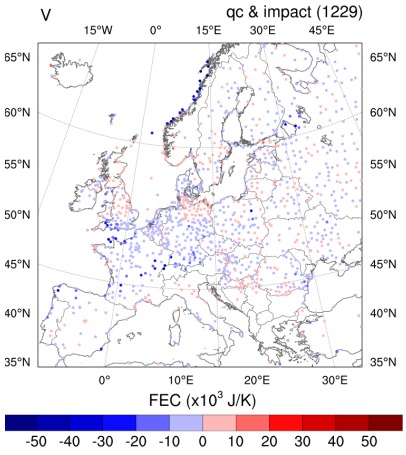
(a) SYNOP 03.08.



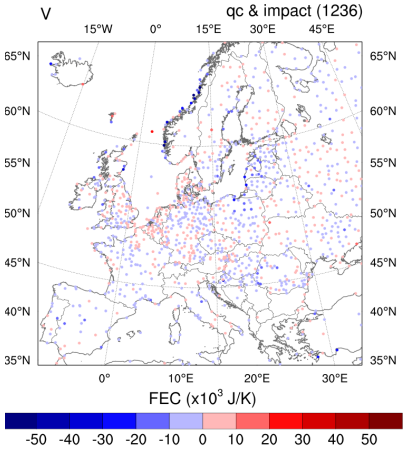
(b) SYNOP 04.08.



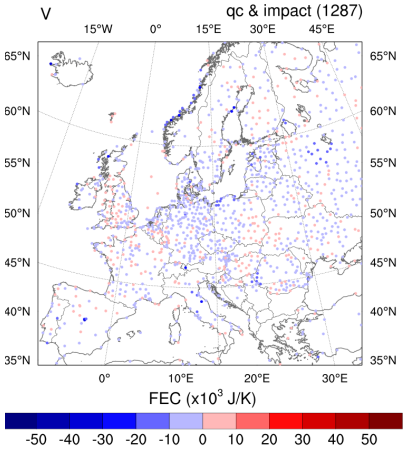
(c) SYNOP 09.08.



(d) SYNOP 10.08.

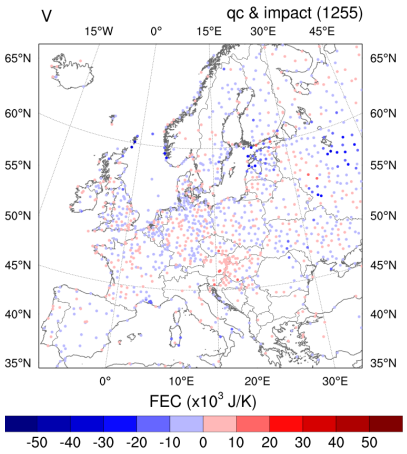


(e) SYNOP 11.08.

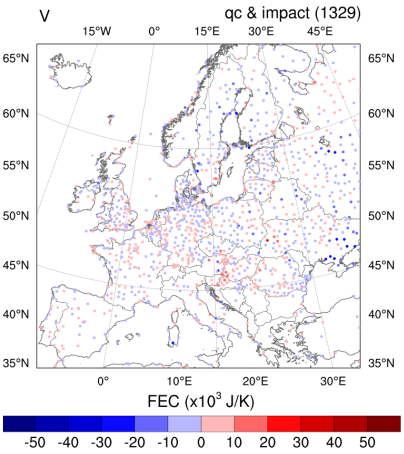


(f) SYNOP 13.08.

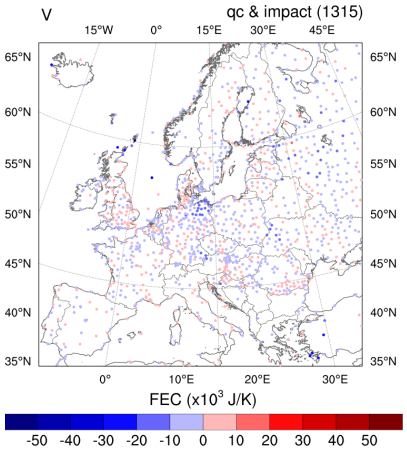
Figure A.3: Impact of SYNOP stations measuring wind (assimilation of meridional V) on the 6-hour forecast, for assimilation hour 06 UTC and for all poor predictability cases.



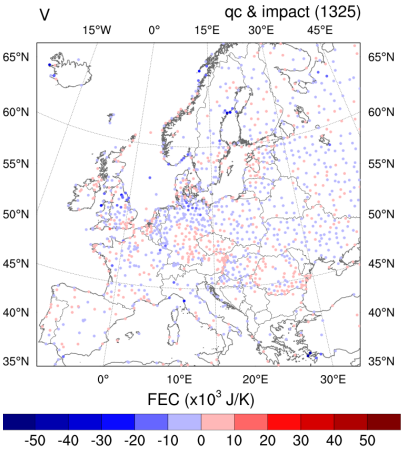
(a) SYNOP 03.08.



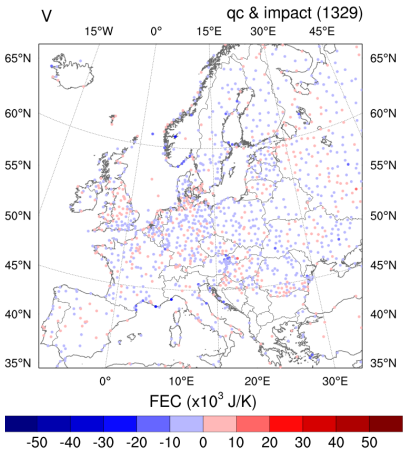
(b) SYNOP 04.08.



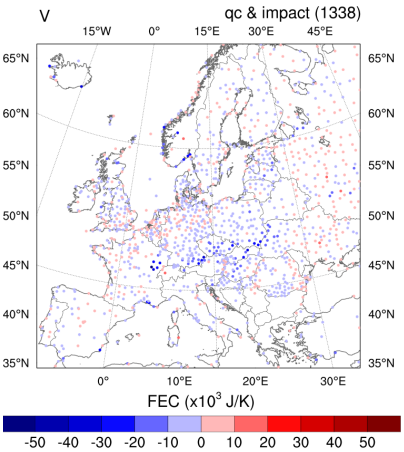
(c) SYNOP 09.08.



(d) SYNOP 10.08.

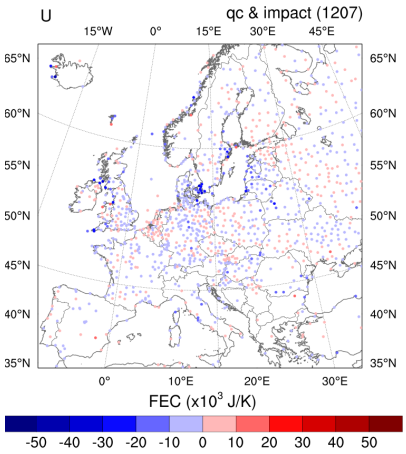


(e) SYNOP 11.08.

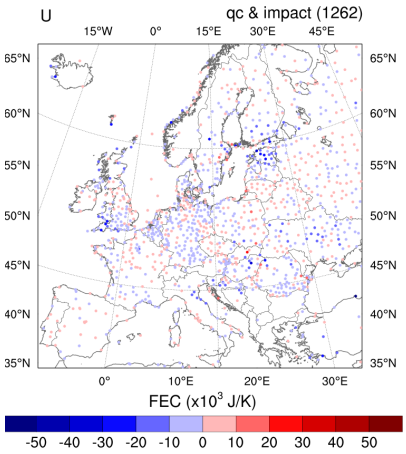


(f) SYNOP 13.08.

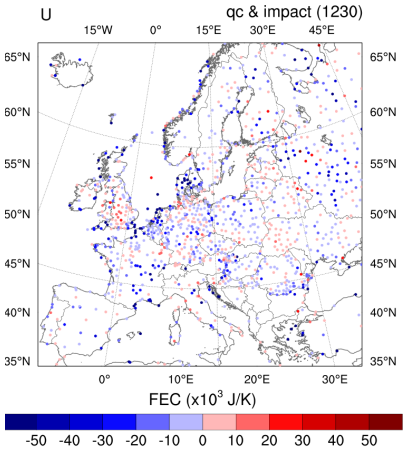
Figure A.4: Impact of SYNOP stations measuring wind (assimilation of meridional V) on the 6-hour forecast, for assimilation hour 12 UTC and for all poor predictability cases.



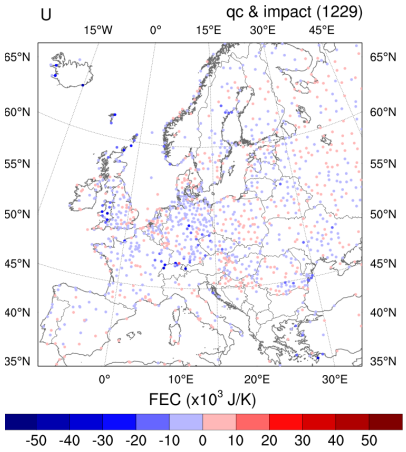
(a) SYNOP 03.08.



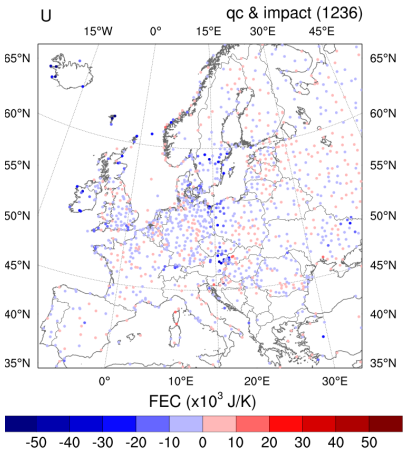
(b) SYNOP 04.08.



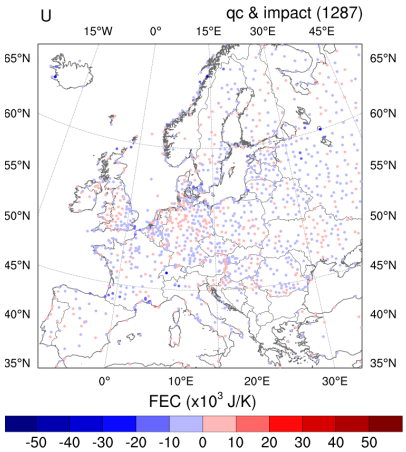
(c) SYNOP 09.08.



(d) SYNOP 10.08.

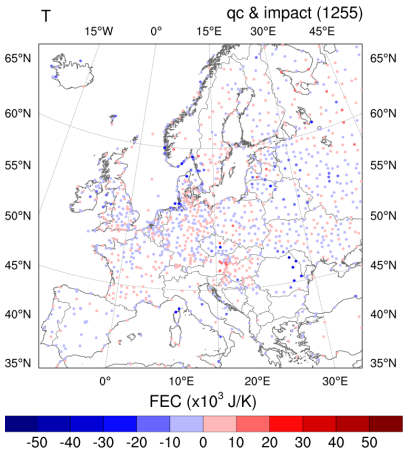


(e) SYNOP 11.08.

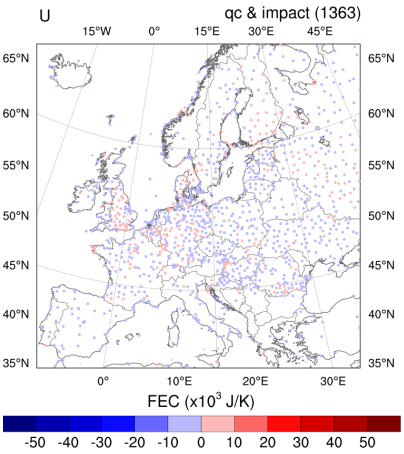


(f) SYNOP 13.08.

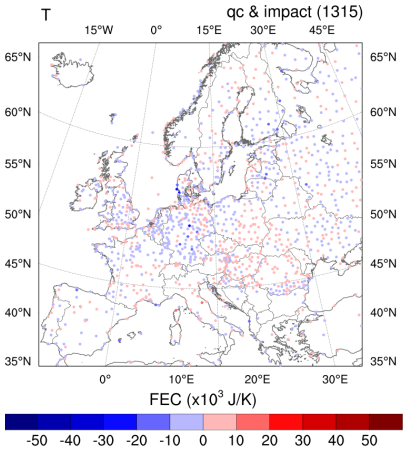
Figure A.5: Impact of SYNOP stations measuring wind (assimilation of zonal U) on the 6-hour forecast, for assimilation hour 06 UTC and for all poor predictability cases.



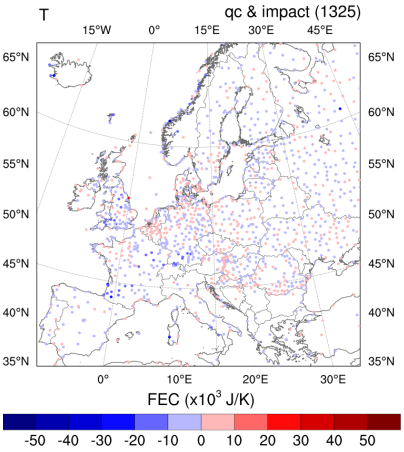
(a) SYNOP 03.08.



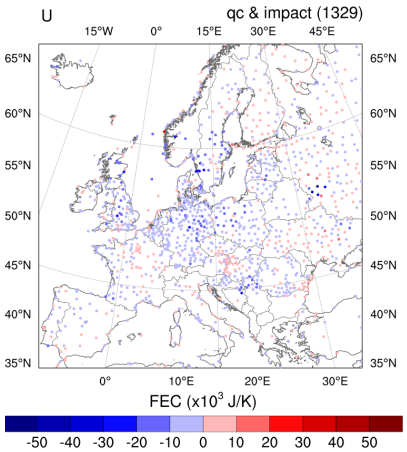
(b) SYNOP 04.08.



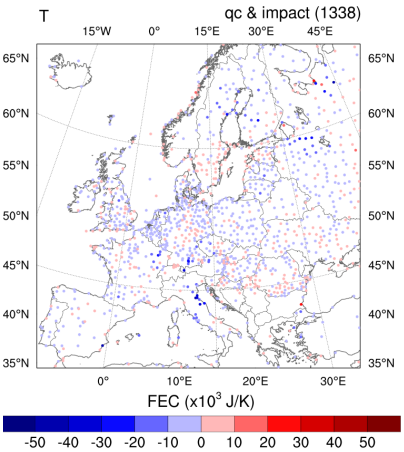
(c) SYNOP 09.08.



(d) SYNOP 10.08.



(e) SYNOP 11.08.



(f) SYNOP 13.08.

Figure A.6: Impact of SYNOP stations measuring wind (assimilation of zonal U) on the 6-hour forecast, for assimilation hour 12 UTC and for all poor predictability cases.

A.3 Model namelist configuration

Table A.1: WRF namelist for the forward runs.

$\&time\_control$
$run\_hours = 06,$
$start\_year = 2014,$
$start\_month = 08,$
$start\_day = 01,$
$start\_hour = 00,$
$start\_minute = 0,$
$start\_second = 0,$
$end\_year = 2014,$
$end\_month = 08,$
$end\_day = 01,$
$end\_hour = 06,$
$end\_minute = 0,$
$end\_second = 0,$
$interval\_seconds = 21600,$
$input\_from\_file = true,$
$input\_outname = "wrfinput\_d < domain > \_ < date > ",$
$write\_input = false,$
$history\_interval = 360,$
$frames\_per\_outfile = 1,$
$debug\_level = 0,$
$inputout\_interval = 360,$
$inputout\_begin\_h = 0,$
$inputout\_end\_h = 06,$
/
$\&domains$
$time\_step = 90,$
$max\_dom = 1,$
$e\_we = 250,$
$e\_sn = 250,$
$e\_vert = 41,$
$num\_metgrid\_levels = 27,$
$p\_top\_requested = 1000,$
$dx = 15000,$
$dy = 15000,$
$parent\_id = 1,$
$i\_parent\_start = 1,$
$j\_parent\_start = 1,$
$smooth\_option = 1,$
/
$\&physics$
$mp\_physics = 4,$
$ra\_lw\_physics = 1,$
$ra\_sw\_physics = 1,$

$radt = 15,$
$sf\_sfclay\_physics = 1,$
$sf\_surface\_physics = 1,$
$bl\_pbl\_physics = 1,$
$cu\_physics = 5,$
$cudt = 5,$
$num\_soil\_layers = 4,$
$num\_land\_cat = 24,$
$mp\_zero\_out = 2,$
/
&dynamics
$dyn\_opt = 2,$
$w\_damping = 0,$
$diff\_opt = 0,$
$km\_opt = 1,$
$dampcoef = 0.2,$
$time\_step\_sound = 6,$
$base\_temp = 290.0,$
/
&bdy_control
$specified = true,$
/
&perturbation
$trajectory\_io = true$

Table A.2: WRFVAR namelist for the assimilation runs.

&wrfvar1
$var4d = false,$
$multi\_inc = 0,$
$print\_detail\_obs = false,$
/
&wrfvar3
$ob\_format = 1,$
$num\_fgat\_time = 1,$
/
&wrfvar4
$thin\_rainobs = false,$
$use\_synopobs = true,$
$use\_shipsobs = true,$
$use\_metarobs = true,$
$use\_soundobs = true,$
$use\_mtgirsobs = false,$
$use\_tamdarobs = false,$
$use\_pilotobs = true,$

$use\_airepobs = false,$
$use\_geoamvobs = false,$
$use\_polaramvobs = false,$
$use\_bogusobs = false,$
$use\_buoyobs = true,$
$use\_profilerobs = false,$
$use\_satemobs = false,$
$use\_gpsztdobs = false,$
$use\_gpspwobs = true,$
$use\_gpsrefobs = false,$
$use\_qscatobs = true,$
$use\_mhsobs = false,$
$use\_amsuaobs = false,$
$use\_amsubobs = false,$
$use\_airsobs = false,$
$use\_airsretobs = false,$
$use\_iasiobs = true,$
/
&$wrfvar5$
$check\_max\_iv = false,$
$put\_rand\_seed = false,$
/
&$wrfvar6$
$ntmax = 50,$
$orthonorm_gradient = true,$
/
&$wrfvar7$
$cv\_options = 5,$
/
&$wrfvar11$
$cv\_options\_hum = 1,$
$check\_rh = 2,$
$seed\_array1 = 2003010100,$
$seed\_array2 = 2003010100,$
$calculate\_cg\_cost\_fn = true,$
/
&$wrfvar12$
/
&$wrfvar13$
/
&$wrfvar14$
$rtminit\_nsensor = 1,$
$rtminit\_platform = 10,$
$rtminit\_satid = 2,$
$rtminit\_sensor = 16,$

```
write_jacobian = false,
qc_rad = true,
write_iv_rad_ascii = false,
write_oa_rad_ascii = true,
rtm_option = 2,
use_crtm_kmatrix = true,
crtm_cloud = false,
only_sea_rad = false,
use_varbc = true,
varbc_factor = 2.0,
crtm_coef_path = "./var/run/crtm_coeffs_2.1.3",
crtm_irland_coef = "USGS.IRland.EmisCoeff.bin",
/
&wrfvar17
analysis_type = "3D − VAR",
adj_sens = false,
/
&wrfvar18
analysis_date = "2014 − 08 − 01_00 : 00 : 00.0000",
/
&wrfvar21
time_window_min = "2014 − 07 − 31_23 : 00 : 00.0000",
/
&wrfvar22
time_window_max = "2014 − 08 − 01_01 : 00 : 00.0000",
/
&perturbation jcdfi_use = false,
trajectory_io = true
```

Table A.3: WRFPLUS namelist for the adjoint runs.

```
&time_control
input_from_file = true,
write_input = false,
io_form_auxhist7 = 2,
restart = false,
debug_level = 0,
iofields_filename = "./Build_WRF3.8/WRFDA/var/run/plus.io_config",
ignore_iofields_warning = true,
/
&physics
mp_physics = 98,
ra_lw_physics = 0,
ra_sw_physics = 0,
radt = 15,
```

$sf\_sfclay\_physics = 0,$
$sf\_surface\_physics = 1,$
$bl\_pbl\_physics = 98,$
$cu\_physics = 0,$
$cudt = 5,$
$num\_soil\_layers = 4,$
$num\_land\_cat = 24,$
$mp\_zero\_out = 2,$
/
&dynamics
$dyn\_opt = 302,$
$w\_damping = 0,$
$diff\_opt = 0,$
$km\_opt = 1,$
$dampcoef = 0.2,$
$time\_step\_sound = 6,$
$base\_temp = 290.0,$
/
&bdy_control
$specified = true,$
/
&perturbation
$trajectory\_io = true,$
$jcdfi\_use = false$

Table A.4: WRFVAR namelist for impact calculation runs.

&time_control
$input\_from\_file = true,$
$input\_outname =$ "$wrfinput\_d < domain > \_ < date >$ ",
$auxinput17\_inname =$ "$./gr01$",
$io\_form\_auxinput17 = 2,$
$history\_interval = 720,$
$frames\_per\_outfile = 1,$
$debug\_level = 0,$
$inputout\_interval = 360,$
$inputout\_begin\_h = 06,$
$inputout\_end\_h = 12,$
$iofields\_filename =$ "$./Build\_WRF3.8/WRFDA/var/run/fso.io\_config$",
/
&physics
$mp\_physics = 4,$
$ra\_lw\_physics = 1,$
$ra\_sw\_physics = 1,$
$radt = 15,$

$sf\_sfclay\_physics = 1,$
$sf\_surface\_physics = 1,$
$bl\_pbl\_physics = 1,$
$cu\_physics = 5,$
$cudt = 5,$
$num\_soil\_layers = 4,$
$num\_land\_cat = 24,$
$mp\_zero\_out = 2,$
$co2tf = 0,$
/
$\&dynamics$
$w\_damping = 0,$
$diff\_opt = 0,$
$km\_opt = 1,$
$dampcoef = 0.2,$
$time\_step\_sound = 6,$
$base\_temp = 290,$
/
$\&bdy\_control,$
$specified = true,$
$real\_data\_init\_type = 3,$
/
$\&wrfvar1$
$var4d = false,$
$multi\_inc = 0,$
$print\_detail\_obs = false,$
/
$\&wrfvar2$
/
$\&wrfvar3$
$ob\_format = 1,$
$num\_fgat\_time = 1,$
/
$\&wrfvar4$
$thin\_rainobs = false,$
$use\_synopobs = true,$
$use\_shipsobs = true,$
$use\_metarobs = true,$
$use\_soundobs = true,$
$use\_mtgirsobs = false,$
$use\_tamdarobs = false,$
$use\_pilotobs = true,$
$use\_airepobs = false,$
$use\_geoamvobs = false,$
$use\_polaramvobs = false,$

$use\_bogusobs = false,$
$use\_buoyobs = true,$
$use\_profilerobs = false,$
$use\_satemobs = false,$
$use\_gpsztdobs = false,$
$use\_gpspwobs = true,$
$use\_gpsrefobs = false,$
$use\_qscatobs = true,$
$use\_mhsobs = false,$
$use\_amsuaobs = false,$
$use\_amsubobs = false,$
$use\_airsobs = false,$
$use\_airsretobs = false,$
$use\_iasiobs = true,$
/
$\&wrfvar5$
$check\_max\_iv = false,$
$put\_rand\_seed = false,$
/
$\&wrfvar6$
$ntmax = 50,$
$eps = 1E - 5,$
$use\_lanczos = true,$
$write\_lanczos = false,$
$orthonorm\_gradient = true,$
/
$\&wrfvar7$
$cv\_options = 5,$ /
$\&wrfvar11$
$cv\_options\_hum = 1,$
$check\_rh = 2,$
$calculate\_cg\_cost\_fn = true,$
/
$\&wrfvar14$
$rtminit\_nsensor = 1,$
$rtminit\_platform = 10,$
$rtminit\_satid = 2,$
$rtminit\_sensor = 16,$
$thinning\_mesh = 120.0,$
$thinning = true,$
$write\_jacobian = false,$
$qc\_rad = true,$
$write\_iv\_rad\_ascii = false,$
$write\_oa\_rad\_ascii = true,$
$rtm\_option = 2,$

$use\_crtm\_kmatrix = true,$
$only\_sea\_rad = false,$
$use\_varbc = true,$
$varbc\_factor = 2.0,$
$crtm\_coef\_path = "./Build\_WRF3.8/WRFDA/var/run/crtm\_coeffs\_2.1.3",$
$crtm\_irland\_coef = "USGS.IRland.EmisCoeff.bin",$
/
$\&wrfvar17$
$analysis\_type = "QC - OBS",$
$sensitivity\_option = 0,$
$adj\_sens = true,$
/
$\&wrfvar18$
$analysis\_date = "2014 - 08 - 01\_00 : 00 : 00.0000",$
/
$\&wrfvar21$
$time\_window\_min = "2014 - 07 - 31\_23 : 00 : 00.0000",$
/
$\&wrfvar22$
$time\_window\_max = "2014 - 08 - 01\_01 : 00 : 00.0000",$
/
$\&perturbation$
$jcdfi\_use = false,$
$trajectory\_io = true$

Bibliography

Ahrens, CD. *Essentials of meteorology: an invitation to the atmosphere*. Cengage Learning, 2011.

Andersson, E. Fourth WMO workshop on the impact of various observing systems on numerical weather prediction wmo, geneva, switzerland 19–21 may 2008. *Abstracts (Session 1)*, 19:21, 2008.

Auligné, T.; McNally, AP., and Dee, DP. Adaptive bias correction for satellite data in a numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 133(624):631–642, 2007.

Baker, N. and Daley, N. Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Q. J. R. Meteorol. Soc.*, 126:1431–1454, 2000.

Barker, DM.; Huang, W.; Guo, YR., and Bourgeois, Al. A three-dimensional variational (3DVAR) data assimilation system for use with MM5. *NCAR Tech Note*, 68, 2003.

Berndt, J. *On the predictability of exceptional error events in wind power forecasting: an ultra large ensemble approach*. PhD thesis, Universität zu Köln, 2018.

Bueno, L. Cornejo; Cuadra, L.; JimÃ©nez-FernÃ¡ndez, S.; Acevedo, J.; Prieto, L., and Salcedo-Sanz, S. Wind power ramp events prediction with hybrid machine learning regression techniques and reanalysis data. *Energies*, 10:1784, 11 2017. doi: 10.3390/en10111784.

Cardinali, C. Monitoring the observation impact on the short range forecast. *Q. J. R. Meteorol. Soc.*, 135:239–250, 2009.

Chalon, G.; Cayla, F., and Diebel, D. Iasi- an advanced sounder for operational meteorology. In *IAF, International Astronautical Congress, 52 nd, Toulouse, France*, 2001.

Clerbaux, C.; Boynard, A.; Clarisse, L.; George, M.; Hadji-Lazaro, J.; Herbin, H.; Hurtmans, D.; Pommier, M.; Razavi, A., and Turquety, S. Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmospheric Chemistry and Physics*, 9 (16):6041–6054, 2009.

Collard, AD. Selection of IASI channels for use in numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 133(629):1977–1991, 2007.

Cui, M.; Zhang, J.; Feng, C.; Florita, AR.; Sun, Y., and Hodge, BM. Characterizing and analyzing ramping events in wind power, solar power, load, and netload. *Renewable Energy*, 111:227–244, 2017.

Daley, R. *Atmospheric data analysis*. Cambridge university press, 1993.

Dennis, K. Data processing at ncep, 2018. URL http://www.emc.ncep.noaa.gov/?data_processing=data_processing.

Dumelow, R. Overview of observing system experiments. In *ECMWF Seminar Proceedings. Seminar on Recent developments in data assimilation for atmosphere and ocean, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, UK*, pages 8–12, 2003.

EUMETSAT, . Iasi level 1:product guide, 2017. URL http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_DMT_151562&RevisionSelectionMethod=LatestReleased&Rendition=Web.

EUMETSAT, . Metosat third generation, 2019. URL https://www.eumetsat.int/website/home/Satellites/FutureSatellites/MeteosatThirdGeneration/index.html.

European Parliament, Council of the European Union. A policy framework for climate and energy in the period from 2020 to 2030, 2014.

Eyre, JR. and Menzel, WP. Retrieval of cloud parameters from satellite sounder data: A simulation study. *Journal of Applied Meteorology*, 28(4):267–275, 1989.

Florita, A.; Hodge, BM., and Orwig, K. Identifying wind and solar ramping events. In *Green Technologies Conference, 2013 IEEE*, pages 147–152. IEEE, 2013.

Freedman, J.; Markus, M., and Penc, R. Analysis of west texas wind plant ramp-up and ramp-down events. *AWS Truewind, LLC*, 2008.

Gambacorta, A. and Barnet, CD. Methodology and information content of the NOAA NESDIS operational channel selection for the Cross-Track Infrared Sounder (CrIS). *IEEE Transactions on Geoscience and Remote Sensing*, 51(6):3207–3216, 2013.

Gelaro, R. and Zhu, Y. Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus A*, 61(2):179–193, 2009.

Gelaro, R.; Buizza, R.; Palmer, TN., and Klinker, E. Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *Journal of the atmospheric sciences*, 55(6):1012–1037, 1998.

Gelaro, R.; Langland, RH.; Rohaly, GD., and Rosmond, TE. An assessment of the singular-vector approach to targeted observing using the FASTEX dataset. *Quarterly Journal of the Royal Meteorological Society*, 125(561):3299–3327, 1999.

Gelaro, R.; Zhu, Y., and Errico, RM. Examination of various-order adjoint-based approximations of observation impact. *Meteorologische Zeitschrift*, 16(6):685–692, 2007.

Goldberg, Mark D; Qu, Yanni; McMillin, Larry M; Wolf, Walter; Zhou, Lihang, and Divakarla, Murty. Airs near-real-time products and algorithms in support of operational numerical weather prediction. *IEEE Transactions on Geoscience and Remote Sensing*, 41(2):379–389, 2003.

Good, GH. Power ensemble evaluation. Technical report, Energy Oriented Center of Excellence for computing applications, 2017.

Guidard, V.; Fourrié, N.; Brousseau, P., and Rabier, F. Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes. *Quarterly Journal of the Royal Meteorological Society*, 137(661):1975–1987, 2011.

Han, Yong. Jcsda community radiative transfer model (crtm): Version 1. 2006.

Hilton, F.; Armante, R.; August, T.; Barnet, C.; Bouchard, A.; Camy-Peyret, C.; Capelle, V.; Clarisse, L.; Clerbaux, C., and Coheur, PF. Hyperspectral Earth observation from IASI: Five years of accomplishments. *bulletin of the american meteorological Society*, 93 (3):347–370, 2012.

Janisková, M. and Cardinali, C. On the impact of the diabatic component in the Forecast Sensitivity Observation Impact diagnostics. In *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III)*, pages 483–511. Springer, 2017.

JCSDA, CRTM. Community Radiative Transfer Model, 2017. URL https://www.jcsda.noaa.gov/projects_crtm.php.

Joo, S.; Eyre, J., and Marriott, R. The impact of MetOp and other satellite data within the Met Office global NWP system using an adjoint-based sensitivity method. *Monthly Weather Review*, 141(10):3331–3342, 2013.

Jülich Supercomputing Centre, . JURECA: Modular supercomputer at Jülich Supercomputing Centre. *Journal of large-scale research facilities*, 4(A132), 2018. doi: 10.17815/jlsrf-4-121-1. URL http://dx.doi.org/10.17815/jlsrf-4-121-1.

Jung, BJ.; Kim, HM.; Auligné, T.; Xin, Z.; Xiaoyan, Z., and Xiang-Yu, H. Adjoint-Derived Observation Impact Using WRF in the Western North Pacific. *Monthly Weather Review*, 141(11):4080–4097, 2013. doi: 10.1175/MWR-D-12-00197.1. URL http://dx.doi.org/10.1175/MWR-D-12-00197.1.

Jung, JA.; Marshall, JF. Le; Riishojgaard, LP., and Derber, JC. The development of hyperspectral infrared water vapor radiance assimilation techniques in the NCEP Global Forecast System. In *ECMWF/EUMETSAT NWP-SAF Workshop on the Assimilation of IASI in NWP*, 2009.

Kalnay, E. *Atmospheric modeling, data assimilation and predictability*. Cambridge university press, 2003.

Kalnay, E.; Ota, Y.; Miyoshi, T., and Liu, J. A simpler formulation of forecast sensitivity to observations: application to ensemble Kalman filters. *Tellus A: Dynamic Meteorology and Oceanography*, 64(1):18462, 2012.

Kleissl, J. *Solar energy forecasting and resource assessment*. Academic Press, 2013.

Langland, RH. and Baker, NL. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus A: Dynamic Meteorology and Oceanography*, 56(3):189–201, 2004.

Langland, RH.; Gelaro, R.; Rohaly, GD., and Shapiro, MA. Targeted observations in FASTEX: Adjoint-based targeting procedures and data impact experiments in IOP17 and IOP18. *Quarterly Journal of the Royal Meteorological Society*, 125(561):3241–3270, 1999.

Liu, Z.; Schwartz, CS.; Snyder, C., and Ha, SY. Impact of assimilating AMSU-A radiances on forecasts of 2008 Atlantic tropical cyclones initialized with a limited-area ensemble Kalman filter. *Monthly Weather Review*, 140(12):4017–4034, 2012.

McNally, AP. A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Quarterly Journal of the Royal Meteorological Society*, 128(585):2551–2556, 2002.

McNally, AP. and Watts, PD. A cloud detection algorithm for high-spectral-resolution infrared sounders. *Quarterly Journal of the Royal Meteorological Society*, 129(595): 3411–3423, 2003.

Metop, Eumetsat. Monitoring weather and climate from space, 2017. URL https://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/index.html.

NCEP, . NCEP ADP Global Upper Air and Surface Weather Observations (PREPBUFR format), may 1997 - continuing, 2008. URL http://rda.ucar.edu/datasets/ds337.0/.

NCEP, . Observations and gridded forcing products input into the NCEP Climate Forecast System Reanalysis. 2015. URL https://doi.org/10.5065/D6HM56HW.

Palmer, TN.; Gelaro, RT.; Barkmeijer, J., and Buizza, R. Singular vectors, metrics, and adaptive observations. *Journal of the Atmospheric Sciences*, 55(4):633–653, 1998.

Parrish, DF. and Derber, JC. The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8):1747–1763, 1992.

Petty, GW. *A First Course in Atmospheric Radiation*. Sundog Pub., 2004. ISBN 9780972903301. URL https://books.google.de/books?id=Q5sRAQAAIAAJ.

Rabier, F.; Klinker, E.; Courtier, Ph., and Hollingsworth, A. Sensitivity of forecast errors to initial conditions. *Quarterly Journal of the Royal Meteorological Society*, 122(529): 121–150, 1996.

Skamarock, WC.; Klemp, JB.; Dudhia, J.; Gill, DO.; Barker, DM.; Duda, MG.; Huang, XY.; Wang, W., and Powers, JG. A Description of the Advanced Research WRF Version 3. *NCAR Tech. Note NCAR/TN-475+STR*, 2008.

TSO-50hertz, . Actual and forecasted total power (GW) values of german wind and solar power plants, 2014. URL https://www.50hertz.com/en/Transparency/GridData/Gridfeed-in.

Van der Vorst, H. Iterative Krylov methods for large linear systems. *Cambridge University Press*, 2003.

Zou, X.; Vandenberghe, F.; Pondeca, M., and Kuo, YH. Introduction to adjoint techniques and the MM5 adjoint modeling system. *NCAR Technical note*, 1997.

General Acknowledgments

The author gratefully acknowledges the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC).

Acknowledgments

I would like to express my gratitude to everyone who supported me during this work. To Dr. Hendrik Elbern, my supervisor, for the guidance and the persistent support during all the phases of the work. I enjoyed our conversations and I will always remember your commitment to your students and the passion for science.
Prof. Dr. Susanne Crewell I thank you for undertaking the co—supervision of the work and I hope you will enjoy reading it. I also would like to thank you for making me feel always welcome in the institute, especially for the last two months of my work.
To all my colleagues in Rhenish Institute for Environmental Research (RIU), it was a pleasure working with you and I hope to cross paths with you again. To Tamari Janalidze, my colleague and friend, thank you for your help, support and friendship. My gratitude goes beyond words. To Dr. Anne C. Lange, for being the kindhearted person she is and for the support from the beginning to the end of the work. To Dr. Philipp Franke, for going in depths with me in the algorithm and for the nice "talk breaks" that made the day better. To Annika Vogel for her thoughtful insights on my work. To Michael and Erna, my roommates in RIU, you made my days brighter.

Last, my love and gratitude to my family who supports all my decisions and gives me courage. Thank you Anna, for being always there for me. A big thank you, to my partner Vasilis for his patience, understanding and support.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit —einschließlich Tabellen, Karten und Abbildungen—, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie —abgesehen von unten angegebenen Teilpublikationen —noch nicht veröffentlicht worden ist, sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von PD Dr. Hendrik Elbern betreut worden.

Datum Paraskevi Vourlioti